Aspects of the Bethe Free Energy

Pascal O. Vontobel
Information Theory Research Group
Hewlett-Packard Laboratories Palo Alto

POA Workshop, Santa Fe, NM, USA, September 2, 2009

Motivation

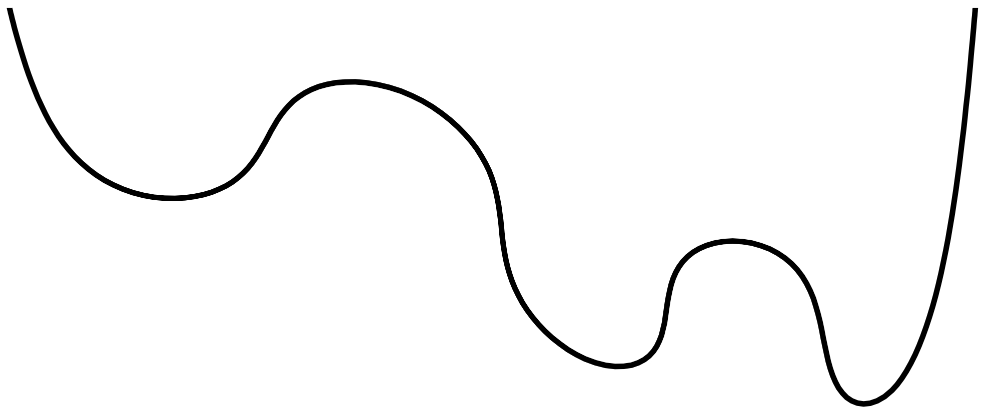
Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the sum-product algorithm (SPA) correspond to stationary points of the Bethe variational free energy (BVFE).

Motivation

Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the sum-product algorithm (SPA) correspond to stationary points of the the Bethe variational free energy (BVFE).

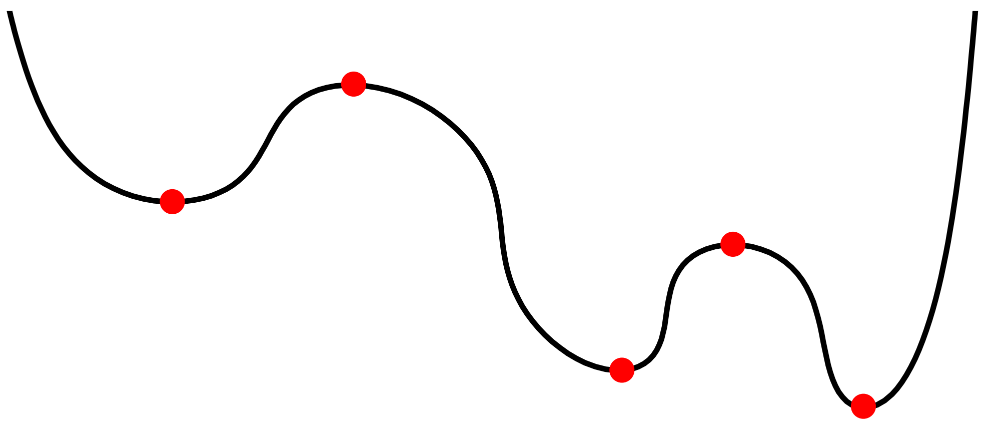


$$F_{\text{Bethe}}(\boldsymbol{\alpha}) = U_{\text{Bethe}}(\boldsymbol{\alpha}) - H_{\text{Bethe}}(\boldsymbol{\alpha}) .$$

Motivation

Theorem (Yedidia/Freeman/Weiss, 2000)

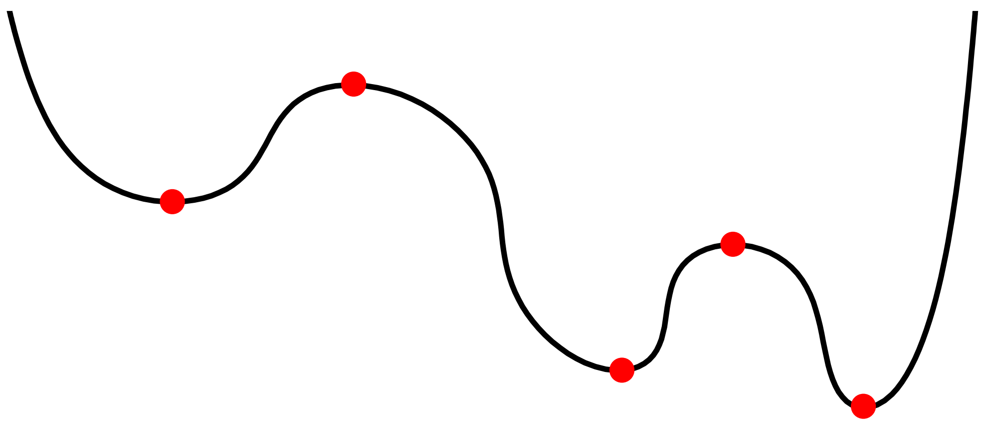
Fixed points of the sum-product algorithm (SPA) correspond to stationary points of the Bethe variational free energy (BVFE).



$$F_{\mathrm{Bethe}}(\boldsymbol{\alpha}) \;=\; U_{\mathrm{Bethe}}(\boldsymbol{\alpha}) \;-\; H_{\mathrm{Bethe}}(\boldsymbol{\alpha}) \;.$$

Motivation

Theorem (Yedidia/Freeman/Weiss, 2000)
Fixed points of the sum-product algorithm (SPA) correspond to stationary points of the Bethe variational free energy (BVFE).

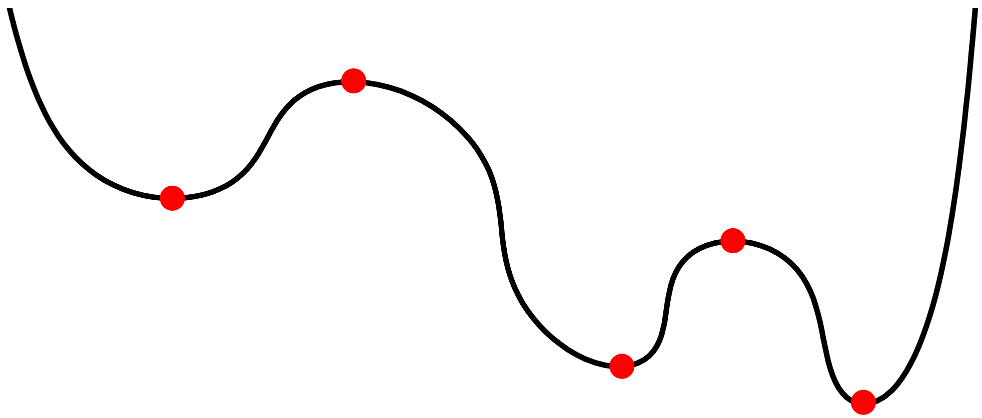


$$F_{\mathrm{Bethe}}(\boldsymbol{\alpha}) = \underbrace{U_{\mathrm{Bethe}}(\boldsymbol{\alpha})}_{\text{linear in } \boldsymbol{\alpha}} - H_{\mathrm{Bethe}}(\boldsymbol{\alpha}) \, .$$

Motivation

Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the sum-product algorithm (SPA) correspond to stationary points of the Bethe variational free energy (BVFE).



$$F_{\mathrm{Bethe}}(\boldsymbol{\alpha}) = \underbrace{U_{\mathrm{Bethe}}(\boldsymbol{\alpha})}_{\text{linear in } \alpha} - \underbrace{H_{\mathrm{Bethe}}(\boldsymbol{\alpha})}_{\text{non-linear in } \alpha} .$$

Overview of Talk

Overview of Talk

- Basics: *codes and graphical models*

Overview of Talk

- Basics: *codes and graphical models*

- Philosophical background of our approach: *graph covers and their relevance for message-passing iterative decoding*

Overview of Talk

- Basics: *codes and graphical models*
- Philosophical background of our approach: *graph covers and their relevance for message-passing iterative decoding*
- Bethe entropy . . .

Overview of Talk

- Basics: *codes and graphical models*
- Philosophical background of our approach: *graph covers and their relevance for message-passing iterative decoding*
- Bethe entropy ...
  - *... and an interpretation of its meaning*

Overview of Talk

- Basics: *codes and graphical models*
- Philosophical background of our approach: *graph covers and their relevance for message-passing iterative decoding*
- Bethe entropy ...
  - *... and an interpretation of its meaning*
  - *... and weight spectra*

Overview of Talk

- Basics: *codes and graphical models*
- Philosophical background of our approach: *graph covers and their relevance for message-passing iterative decoding*
- Bethe entropy . . .
  - *. . . and an interpretation of its meaning*
  - *. . . and weight spectra*
  - *. . . and the edge zeta function*

Graphical representation of a code

Binary Linear Codes

Let $\mathbf{H}$ be a parity-check matrix, e.g.

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Binary Linear Codes

Let $\mathbf{H}$ be a parity-check matrix, e.g.

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

The code $\mathcal{C}$ described by $\mathbf{H}$ is then

$$\mathcal{C} = \left\{ (x_1, x_2, x_3, x_4, x_5) \in \mathbb{F}_2^5 \;\middle|\; \mathbf{H} \cdot \mathbf{x}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}} \pmod 2 \right\}.$$

Binary Linear Codes

Let $\mathbf{H}$ be a parity-check matrix, e.g.

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

The code $\mathcal{C}$ described by $\mathbf{H}$ is then

$$\mathcal{C} = \left\{ (x_1, x_2, x_3, x_4, x_5) \in \mathbb{F}_2^5 \;\middle|\; \mathbf{H} \cdot \mathbf{x}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}} \pmod{2} \right\}.$$

A vector $\mathbf{x} \in \mathbb{F}_2^5$ is a codeword if and only if

$$\mathbf{H} \cdot \mathbf{x}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}} \pmod{2}.$$

Binary Linear Codes

This means that $\mathbf{x}$ is a codeword if and only if $\mathbf{x}$ fulfills the following two equations:

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Binary Linear Codes

This means that $\mathbf{x}$ is a codeword if and only if $\mathbf{x}$ fulfills the following two equations:

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix} \quad \Rightarrow \quad x_1 + x_2 + x_3 = 0 \pmod 2$$

Binary Linear Codes

This means that $\mathbf{x}$ is a codeword if and only if $\mathbf{x}$ fulfills the following two equations:

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{l} x_1 + x_2 + x_3 = 0 \pmod 2 \\ x_2 + x_4 + x_5 = 0 \pmod 2 \end{array}$$

Binary Linear Codes

This means that $\mathbf{x}$ is a codeword if and only if $\mathbf{x}$ fulfills the following two equations:

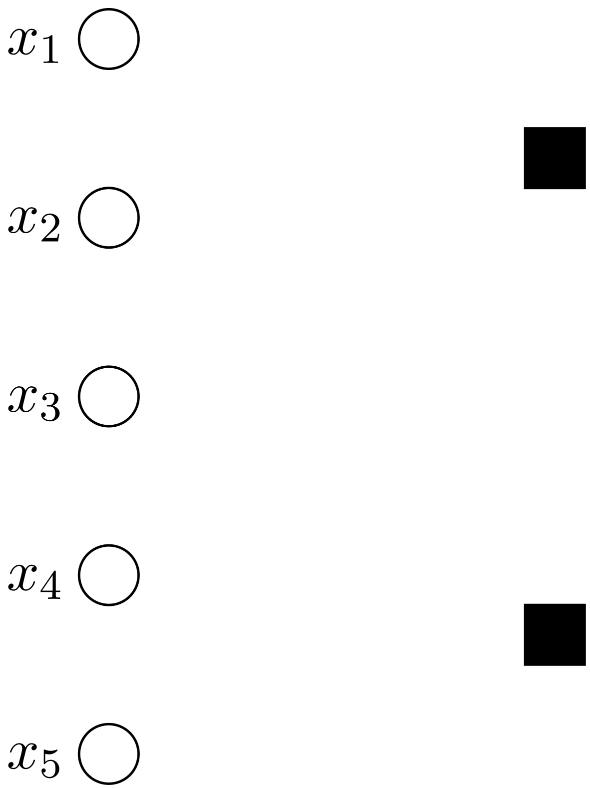
$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{c} x_1 + x_2 + x_3 = 0 \pmod 2 \\ x_2 + x_4 + x_5 = 0 \pmod 2 \end{array}$$

In summary,

$$\begin{aligned} \mathcal{C} &= \left\{ (x_1, x_2, x_3, x_4, x_5) \in \mathbb{F}_2^5 \;\middle|\; \mathbf{H} \cdot \mathbf{x}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}} \text{ (mod 2)} \right\} \\ &= \left\{ (x_1, x_2, x_3, x_4, x_5) \in \mathbb{F}_2^5 \;\middle|\; \begin{array}{c} x_1 + x_2 + x_3 = 0 \text{ (mod 2)} \\ x_2 + x_4 + x_5 = 0 \text{ (mod 2)} \end{array} \right\}. \end{aligned}$$

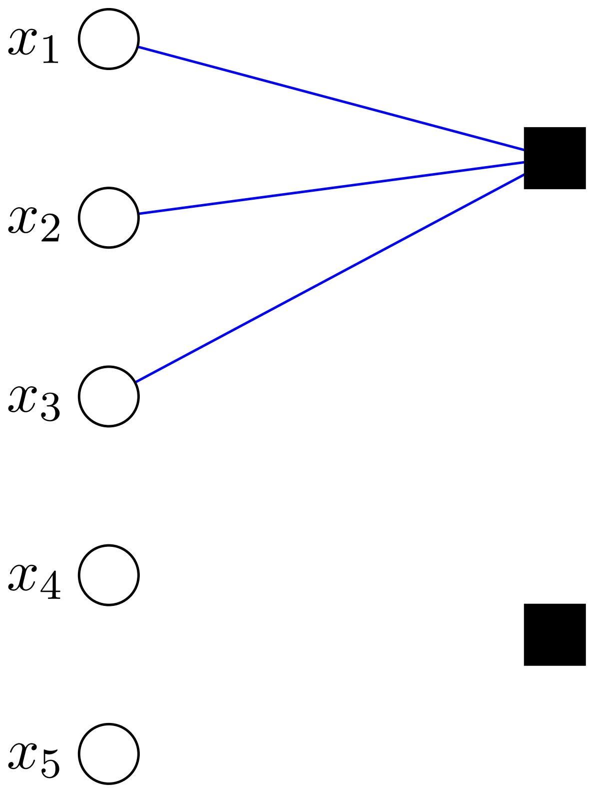
Graphical Representation of a Code

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



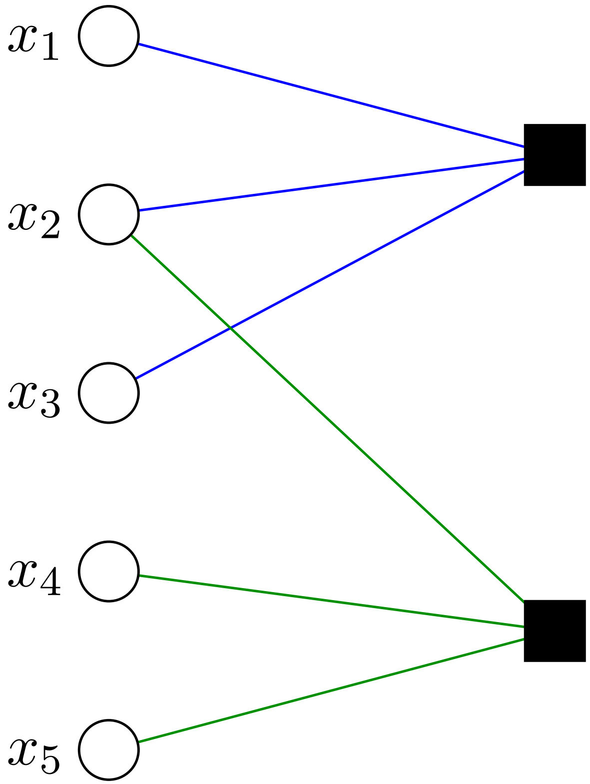
Graphical Representation of a Code

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



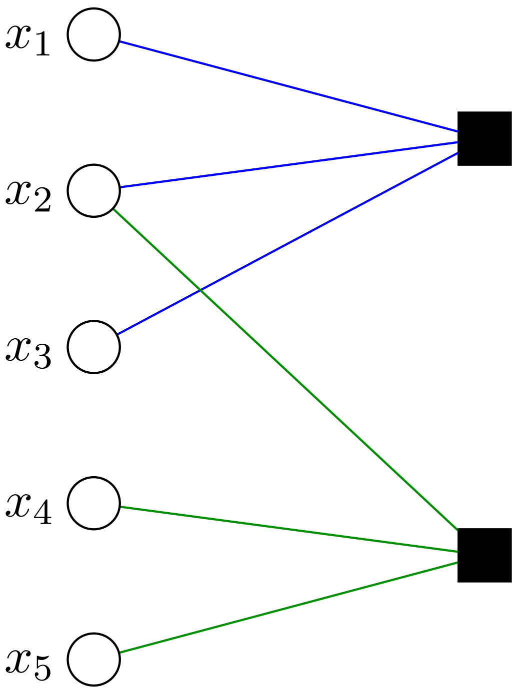
Graphical Representation of a Code

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



FG of a Data Communication System based on a Parity-Check Code

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

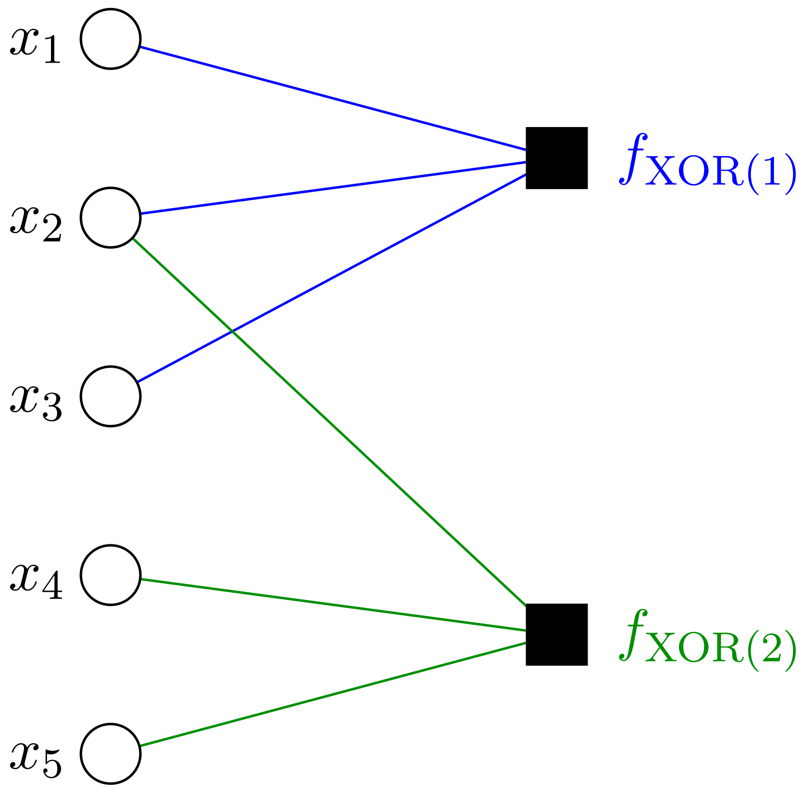


$$f_{\text{XOR}}(x_1, x_2, x_3) \;=\; [x_1 + x_2 + x_3 = 0 \pmod 2]$$

$$f_{\text{XOR}}(x_2, x_4, x_5) \;=\; [x_2 + x_4 + x_5 = 0 \pmod 2]$$

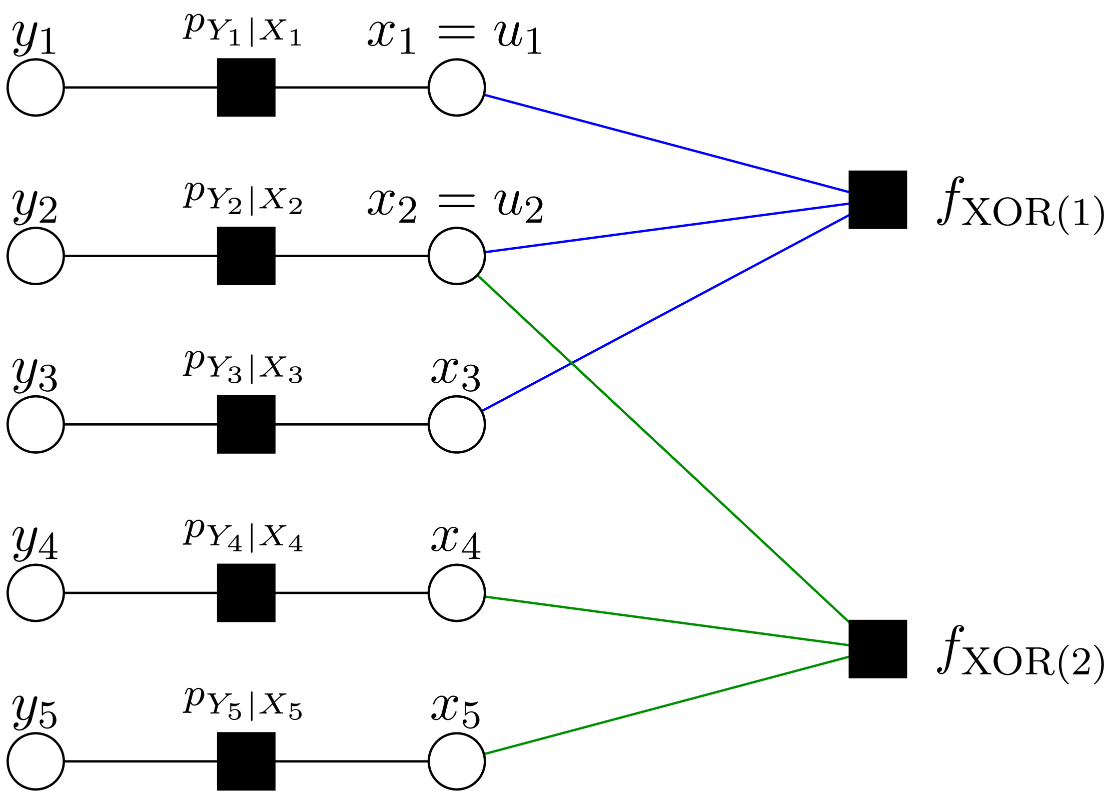
FG of a Data Communication System based on a Parity-Check Code

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



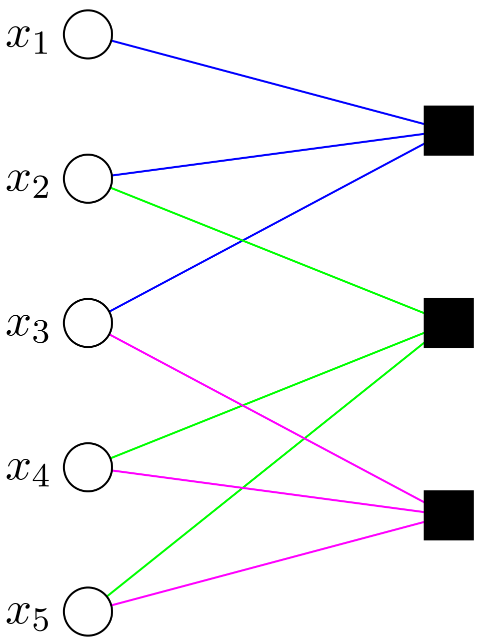
FG of a Data Communication System based on a Parity-Check Code

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



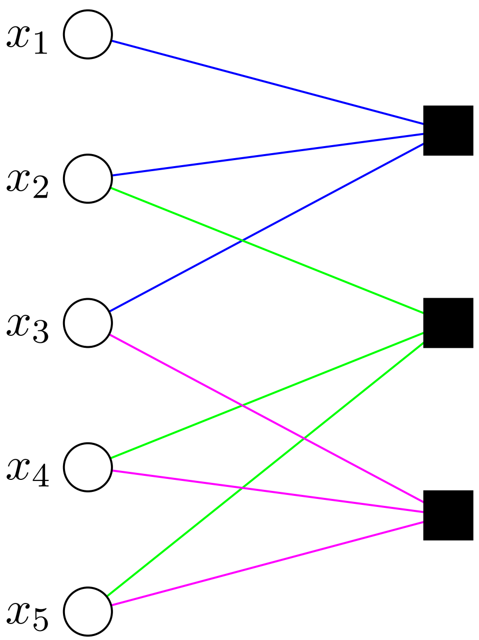
Factor Graph / Tanner Graph of a Binary Linear Code

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$



Factor Graph / Tanner Graph of a Binary Linear Code

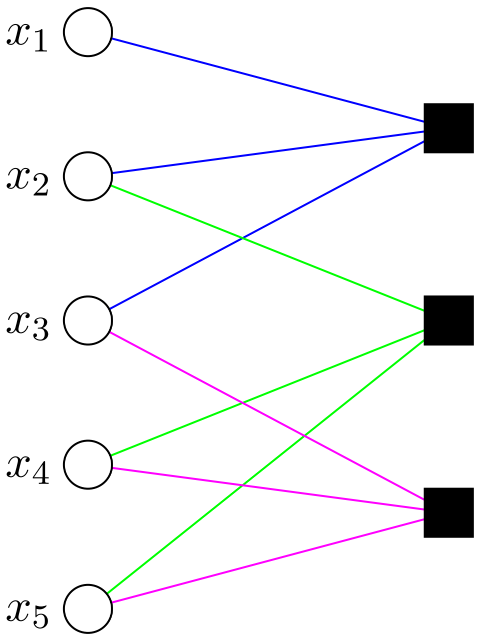
$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$



- Low-density parity-check codes (LDPC) codes are defined by parity-check matrices with very few ones.

Factor Graph / Tanner Graph of a Binary Linear Code

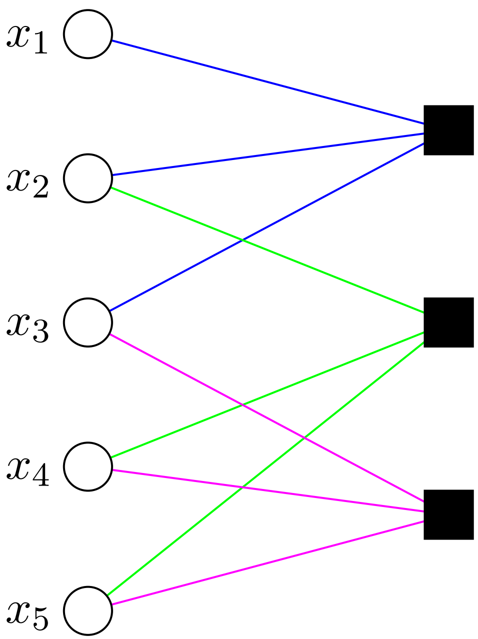
$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$



- Low-density parity-check codes (LDPC) codes are defined by parity-check matrices with very few ones.
- A (j,k)-regular LDPC code is a code whose Tanner graph has uniform variable node degree j and uniform check node degree k.

Factor Graph / Tanner Graph of a Binary Linear Code

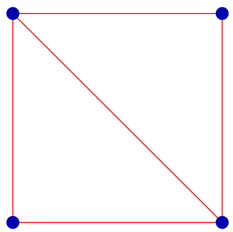
$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$



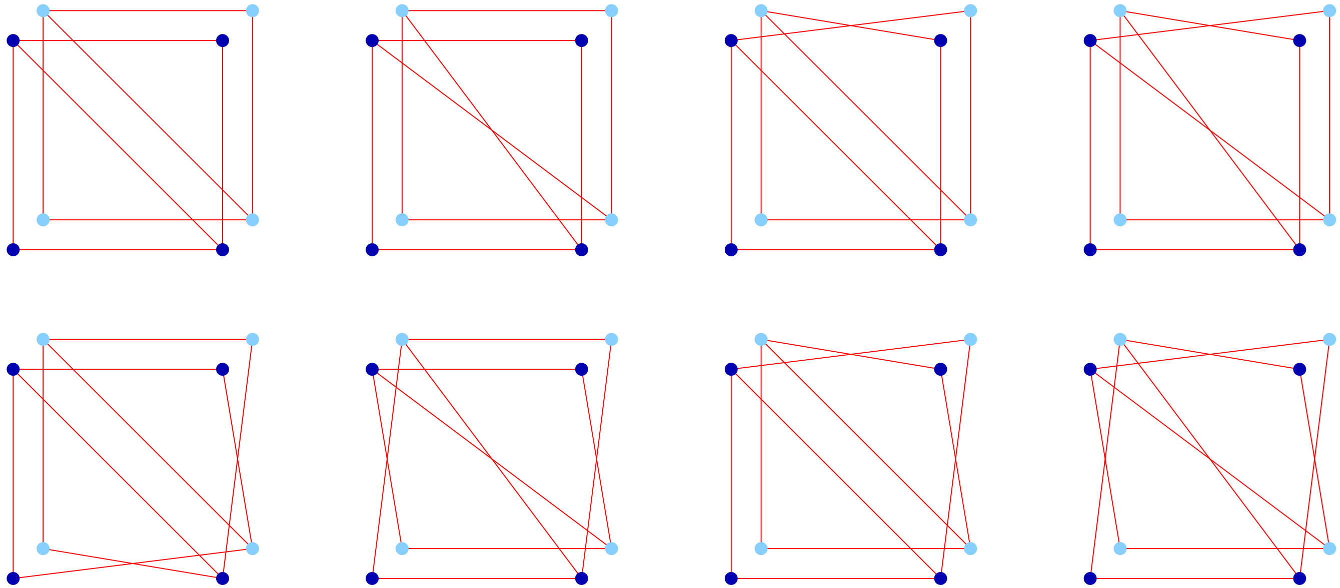
- Low-density parity-check codes (LDPC) codes are defined by parity-check matrices with very few ones.
- A (j, k)-regular LDPC code is a code whose Tanner graph has uniform variable node degree j and uniform check node degree k.
- One can show that Tanner graphs of good codes have cycles. (We assume bounded alphabet size and bounded subcode complexity.)

Graph covers and their relevance for message-passing iterative decoding

Graph Covers



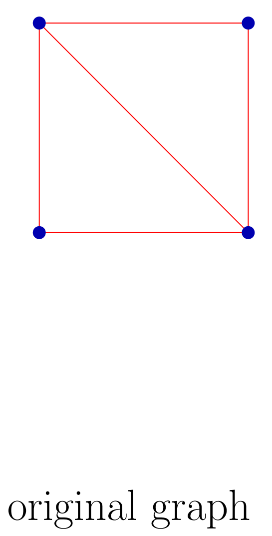
original graph

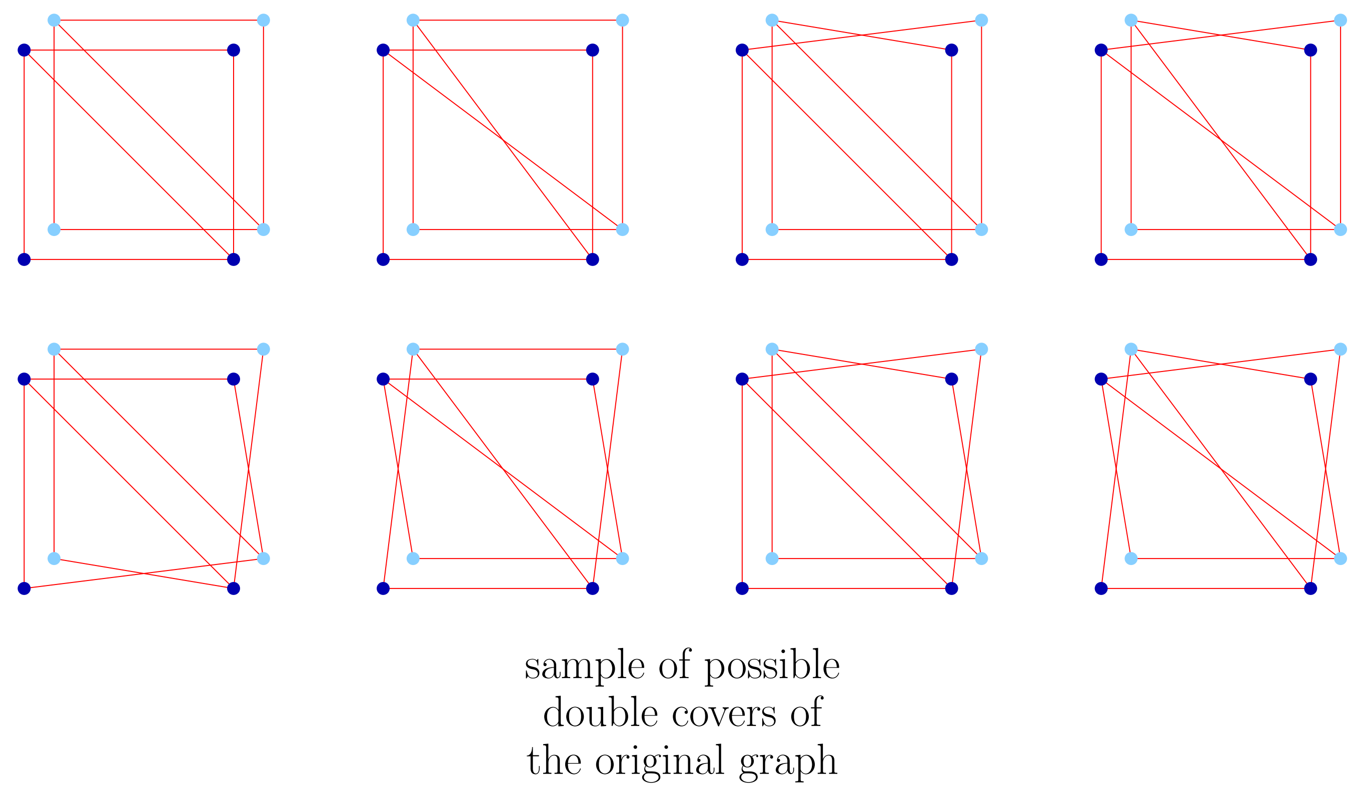


sample of possible
double covers of
the original graph

Definition: A double cover of a graph is . . .

Graph Covers

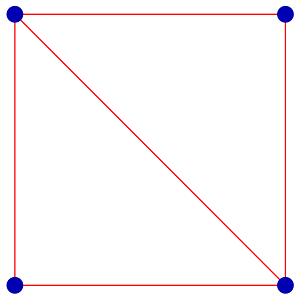




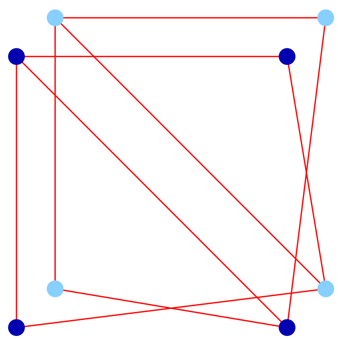
Definition: A double cover of a graph is ...

Note: the above graph has $2! \cdot 2! \cdot 2! \cdot 2! \cdot 2! = 32$ double covers.

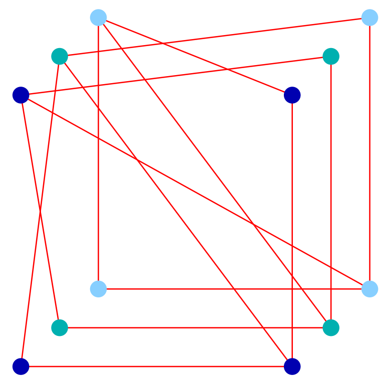
Graph Covers



original graph



(a possible)
double cover of
the original graph

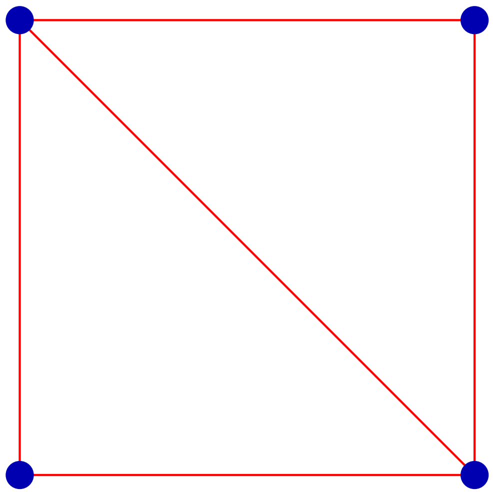


(a possible)
triple cover of
the original graph

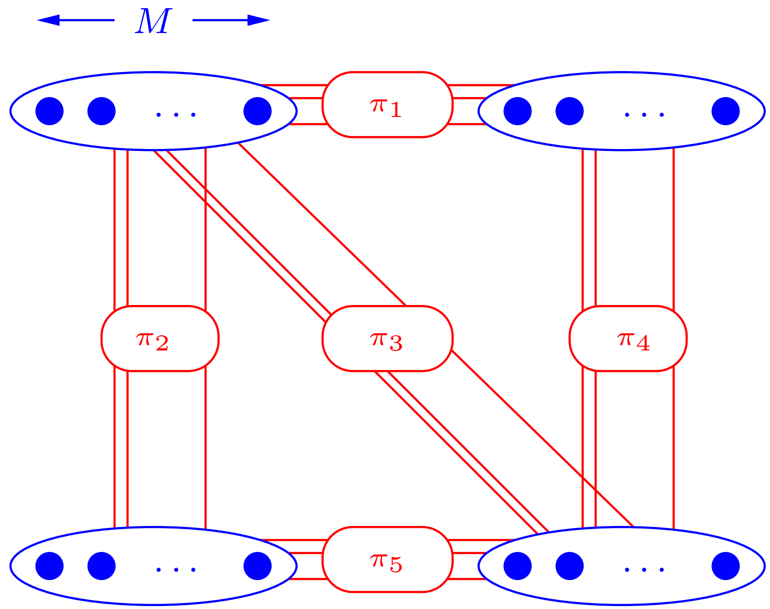
$\ldots$

Besides double covers, a graph also has many triple covers, quadruple covers, quintuple covers, etc.

Graph Covers



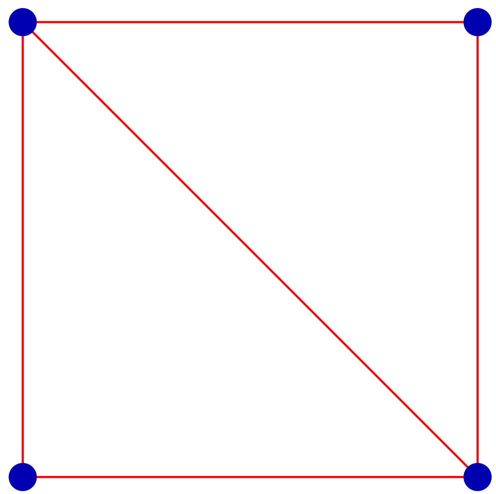
original graph



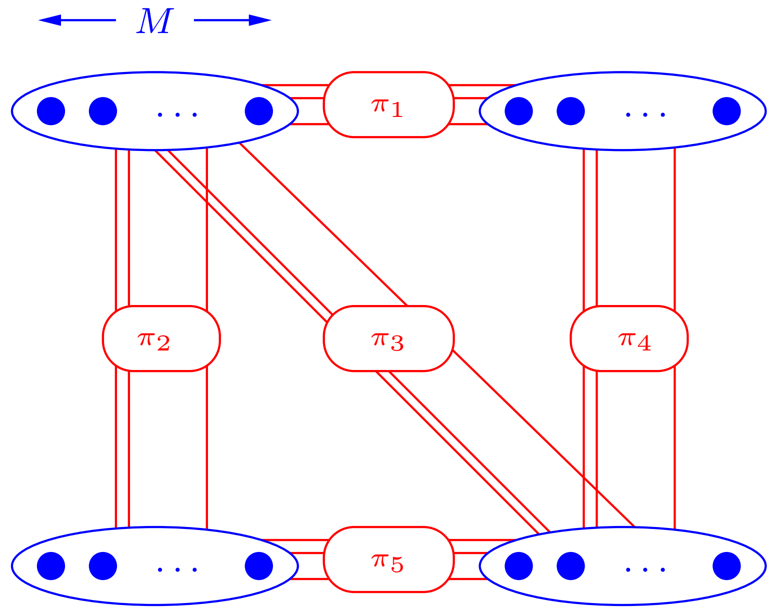
(possible)
M-fold cover of
original graph

An M-fold cover is also called a cover of degree M. Do not confuse this degree with the degree of a vertex!

Graph Covers



original graph



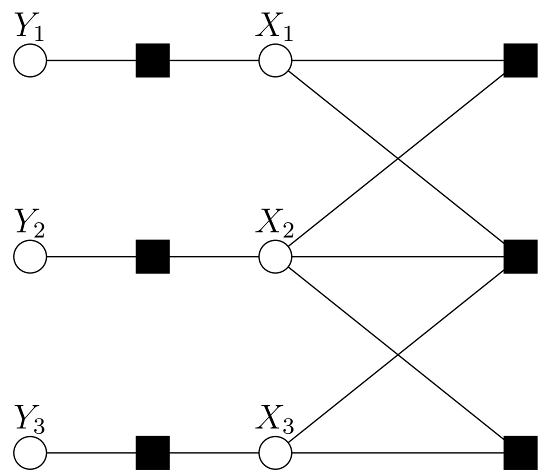
(possible)
M-fold cover of
original graph

An M-fold cover is also called a cover of degree M. Do not confuse this degree with the degree of a vertex!

Note: a graph G with E edges has $(M!)^E$ M-fold covers.

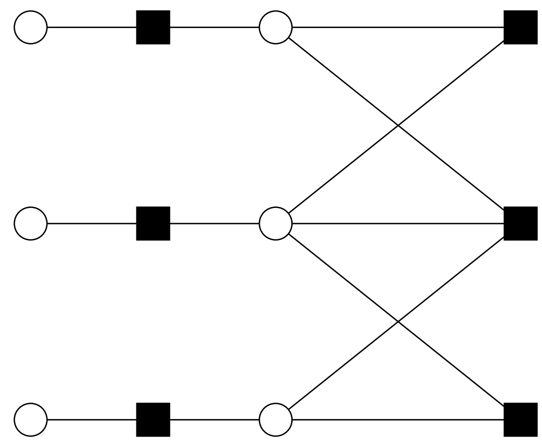
Message-Passing Iterative Decoding

Consider this factor graph:

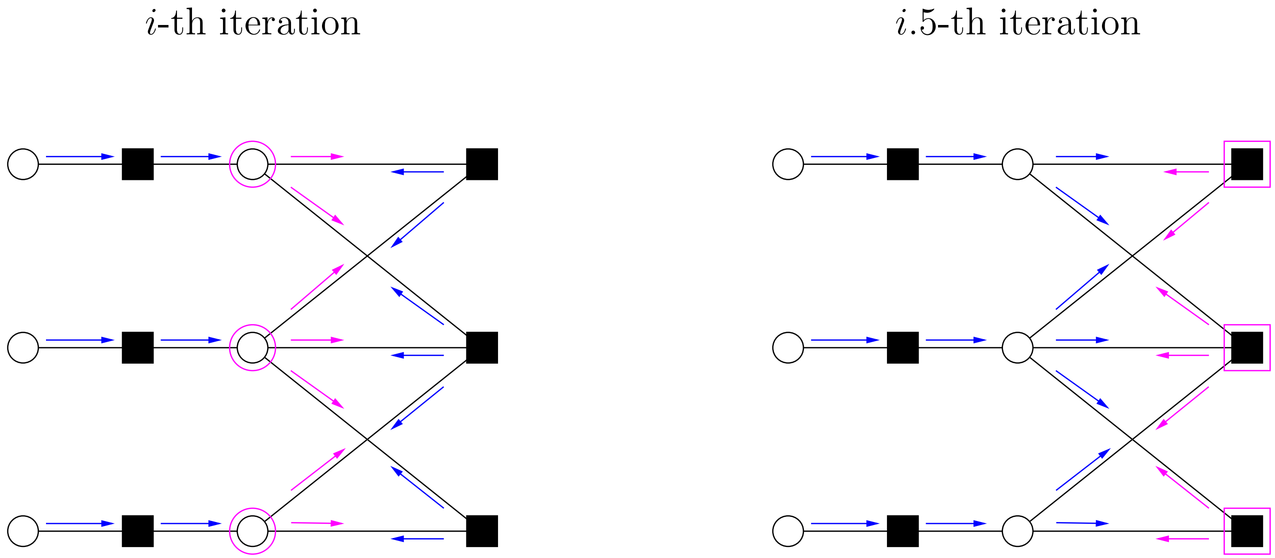


Message-Passing Iterative Decoding

Consider this factor graph:

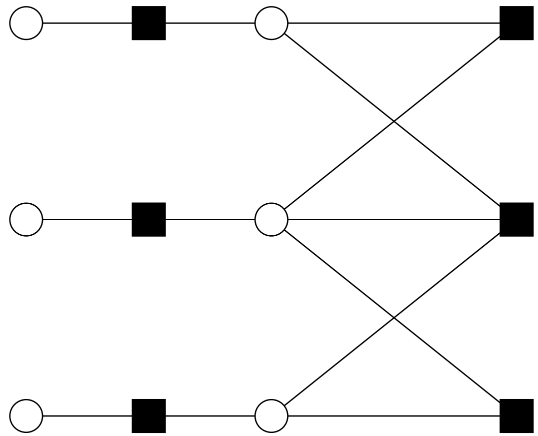


Message-Passing Iterative Decoding



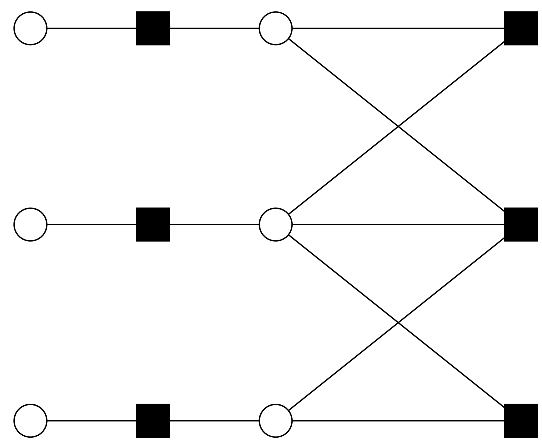
Message-Passing Iterative Decoding

Consider this factor graph:

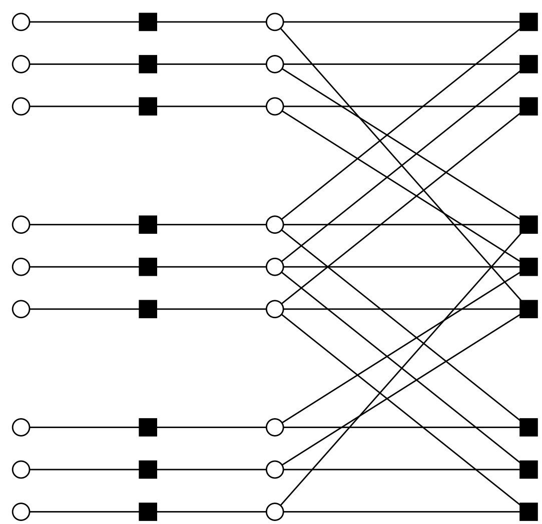


Message-Passing Iterative Decoding

Consider this factor graph:

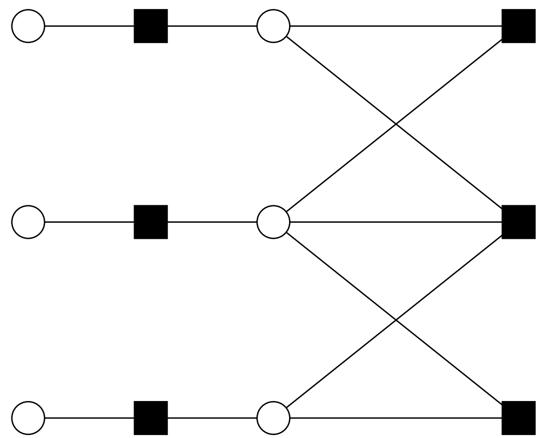


Here is a so-called triple cover of the above factor graph:

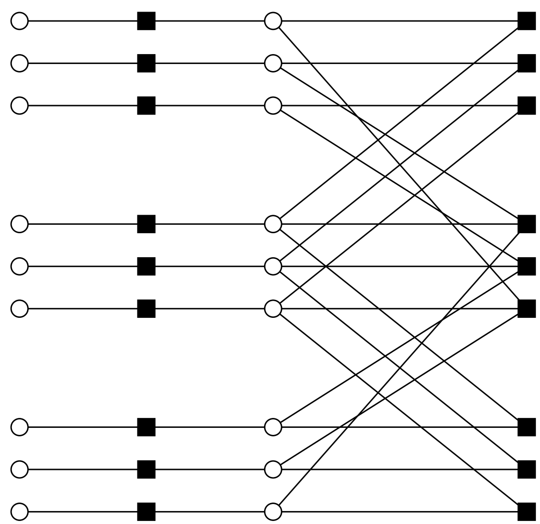


Message-Passing Iterative Decoding

Consider this factor graph:



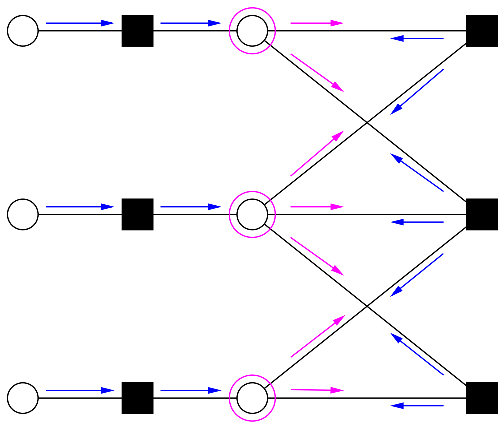
Here is a so-called triple cover of the above factor graph:



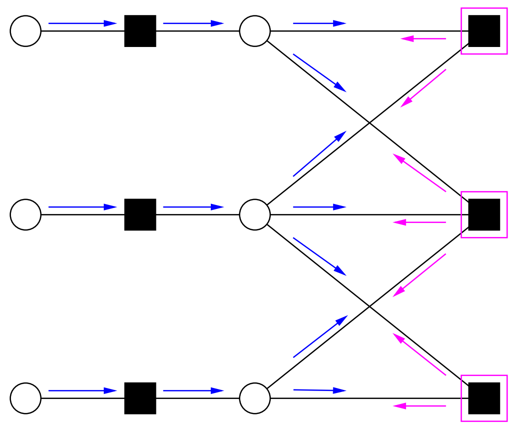
Why do factor graph covers matter for MPI decoding?

Message-Passing Iterative Decoding

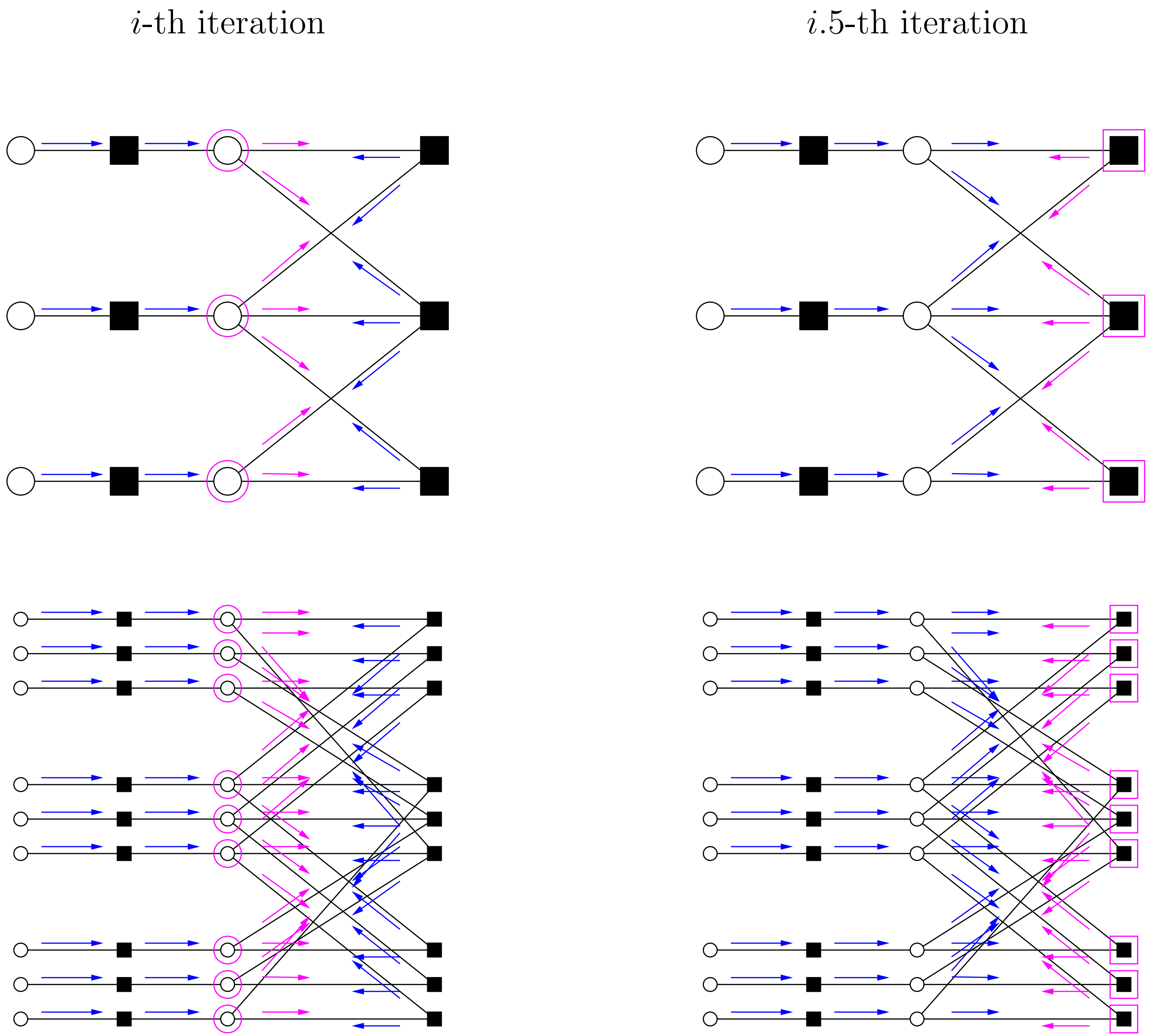
i-th iteration



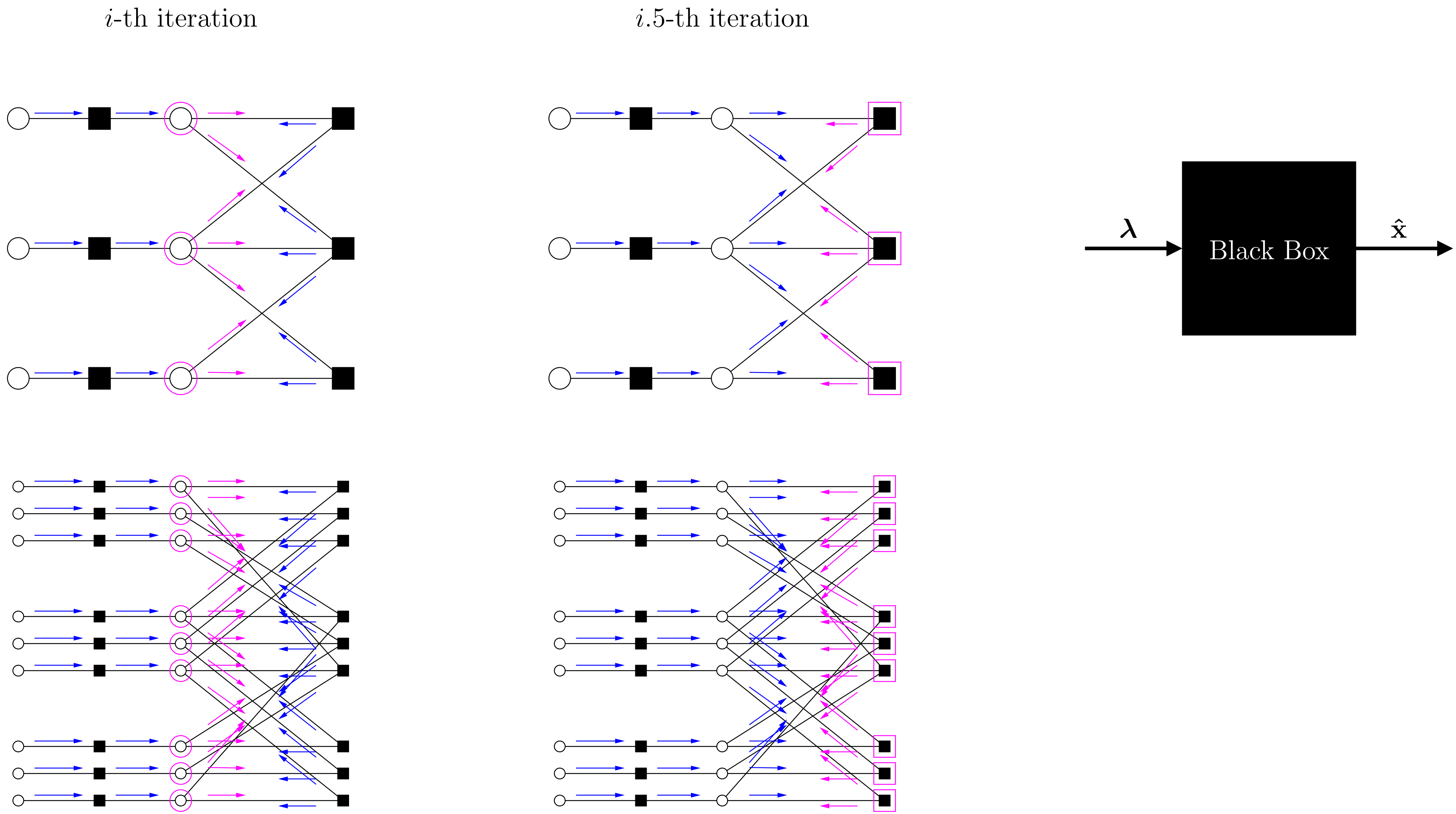
$i.5$-th iteration



Message-Passing Iterative Decoding

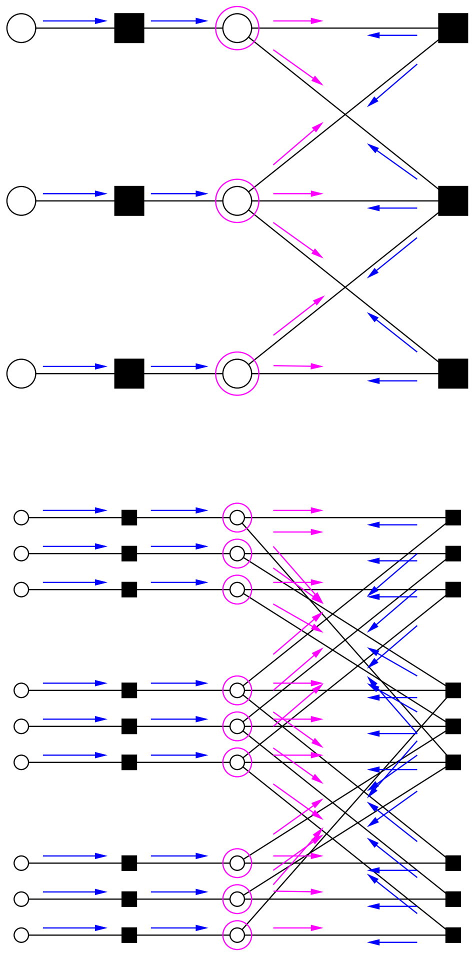


Message-Passing Iterative Decoding

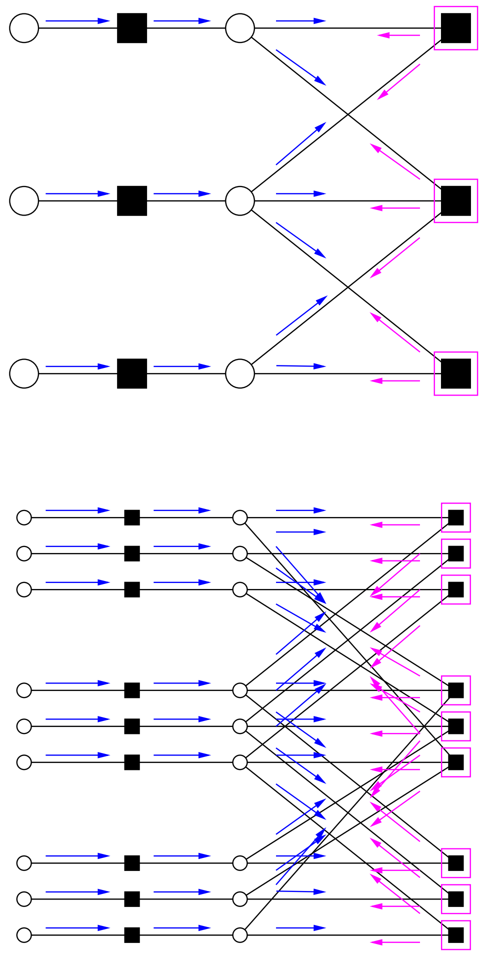


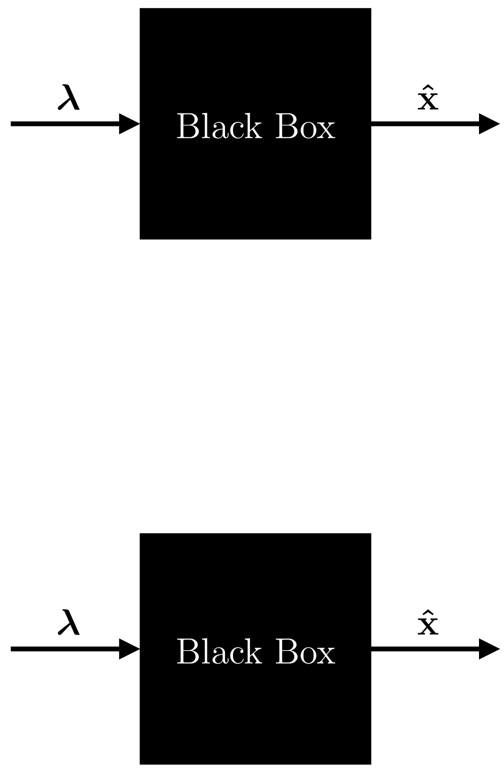
Message-Passing Iterative Decoding

i-th iteration



$i.5$-th iteration





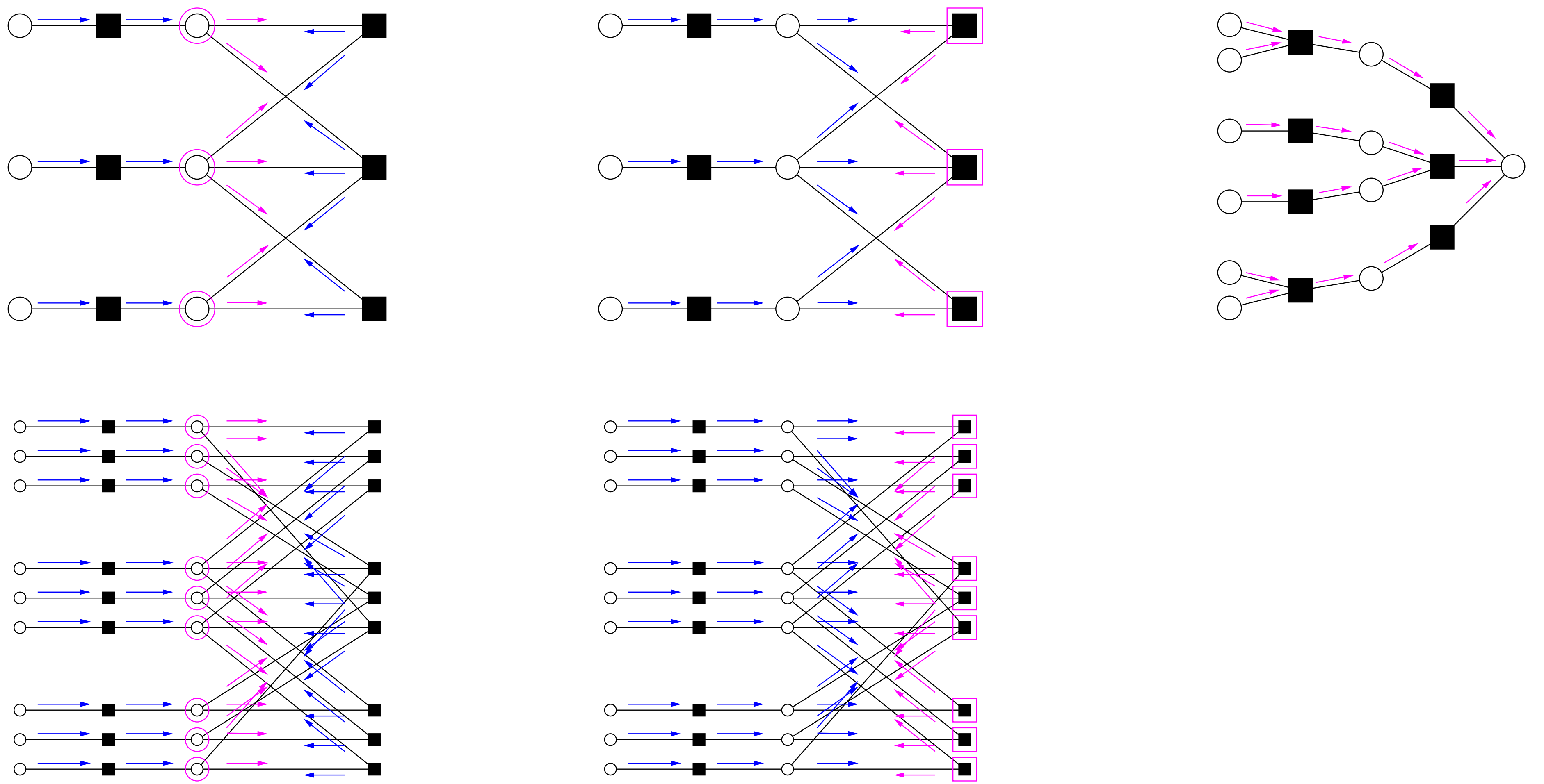
Message-Passing Iterative Decoding

i-th iteration

$i.5$-th iteration

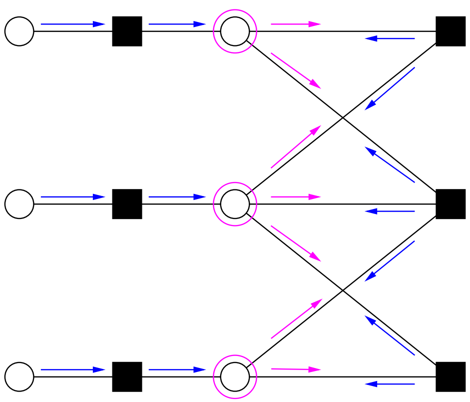
computation tree (without channel function node

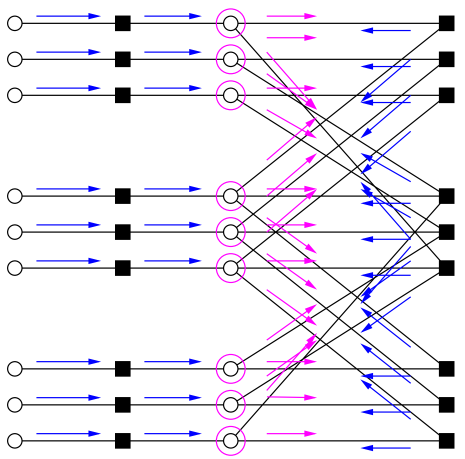
. . . where root is bit node 2



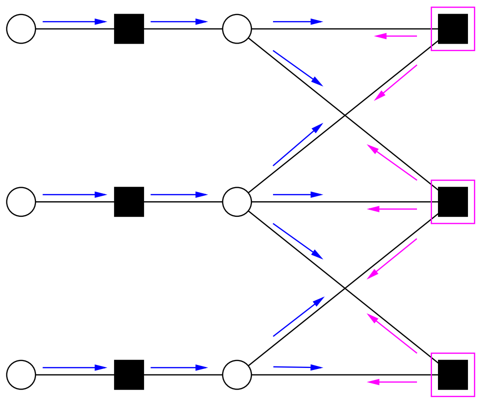
Message-Passing Iterative Decoding

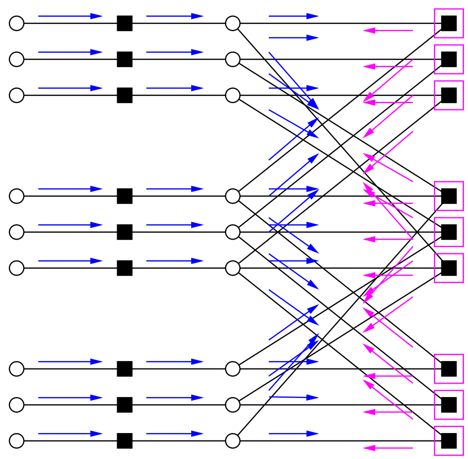
i-th iteration





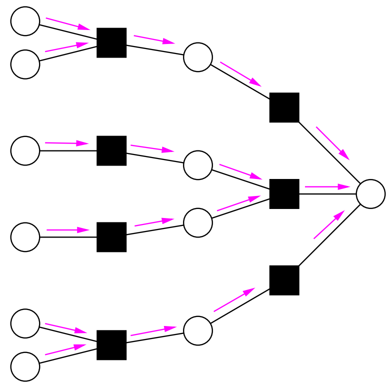
$i.5$-th iteration



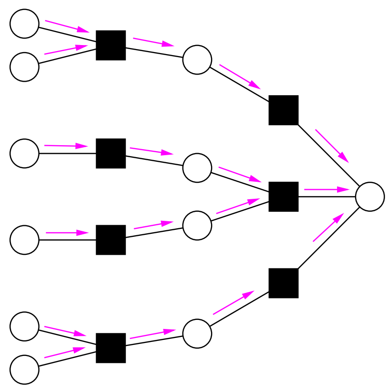


computation tree (without channel function node

. . . where root is bit node 2



. . . where root is a copy of bit node 2



Message-Passing Iterative Decoding

Why do factor graph covers matter?

Message-Passing Iterative Decoding

Why do factor graph covers matter?

Well, a locally operating decoding algorithm cannot distinguish if it is decoding on the original factor graph or on any of its covers.

Message-Passing Iterative Decoding

Why do factor graph covers matter?

Well, a locally operating decoding algorithm cannot distinguish if it is decoding on the original factor graph or on any of its covers.

all codewords from all covers are also competing to be the best!

Message-Passing Iterative Decoding

Three questions:

Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?

Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?
  Yes!

Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?
  Yes!
- Can we characterize the set of codewords given by graph covers?

Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?
  Yes!
- Can we characterize the set of codewords given by graph covers?
  Yes!

Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?
  Yes!

- Can we characterize the set of codewords given by graph covers?
  Yes! $\Rightarrow$ Local marginal polytope, i.e., domain of Bethe entropy

Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?
  Yes!
- Can we characterize the set of codewords given by graph covers?
  Yes! $\Rightarrow$ Local marginal polytope, i.e., domain of Bethe entropy
- Can we somehow count the codewords given by graph covers?

Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?
  Yes!
- Can we characterize the set of codewords given by graph covers?
  Yes! $\Rightarrow$ Local marginal polytope, i.e., domain of Bethe entropy
- Can we somehow count the codewords given by graph covers?
  Yes!

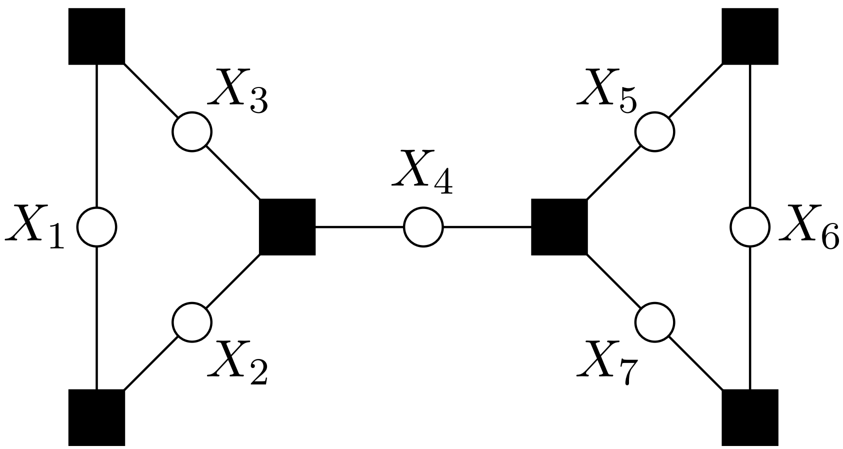
Message-Passing Iterative Decoding

Three questions:

- Are there codewords in graph covers that cannot be explained by codewords in the base graph?
  Yes!
- Can we characterize the set of codewords given by graph covers?
  Yes! $\Rightarrow$ Local marginal polytope, i.e., domain of Bethe entropy
- Can we somehow count the codewords given by graph covers?
  Yes! $\Rightarrow$ Bethe entropy

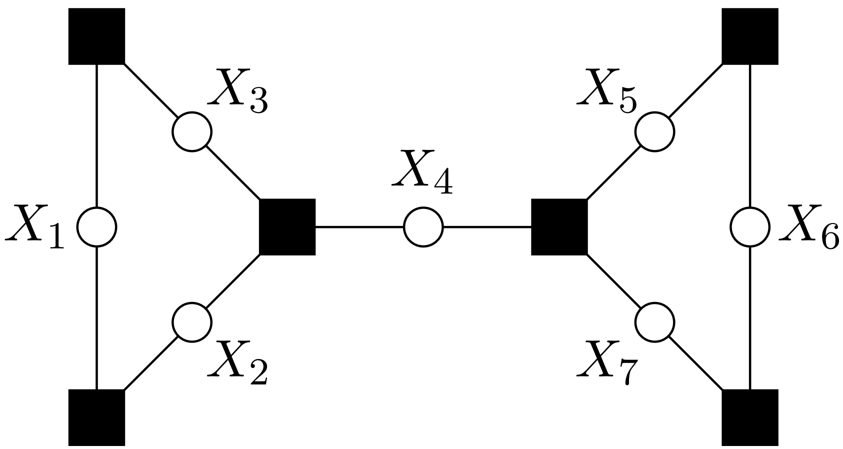
Codewords in graph covers

Codewords in Graph Covers

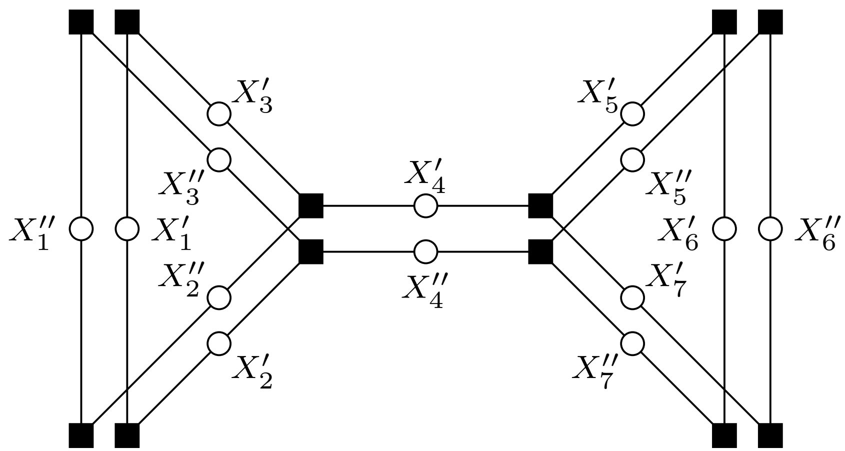


Base factor/Tanner graph
of a length-7 code

Codewords in Graph Covers

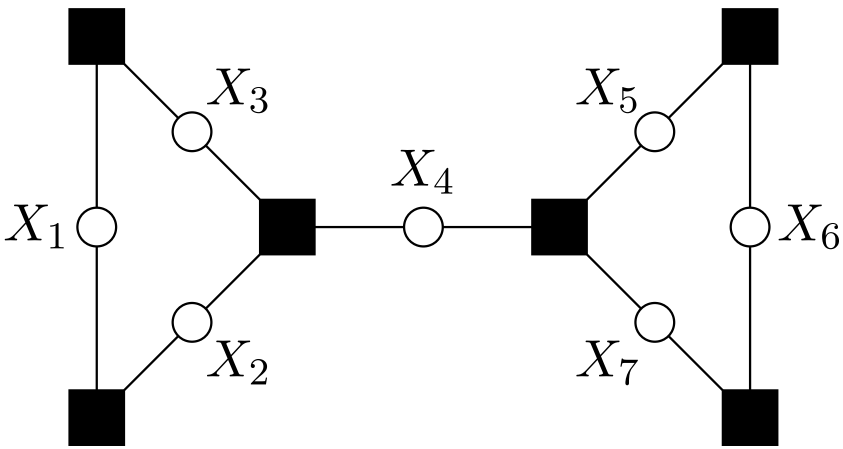


Base factor/Tanner graph
of a length-7 code

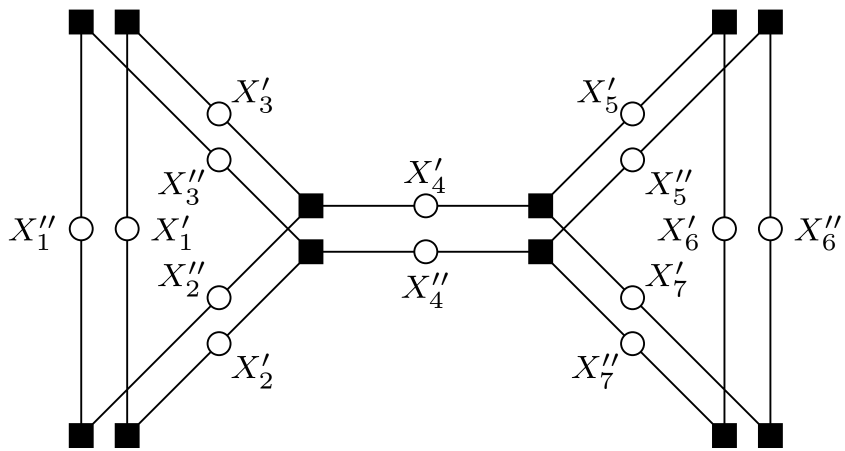


Possible double cover of
the base factor graph

Codewords in Graph Covers



Base factor/Tanner graph of a length-7 code

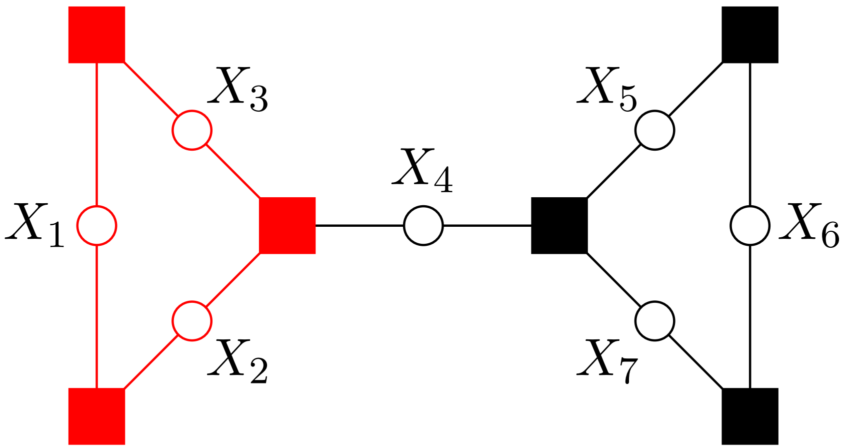


Possible double cover of the base factor graph

Let us study the codes defined by the graph covers of the base Tanner/factor graph.

Codewords in Graph Covers

Obviously, any codeword in the base normal factor graph can be lifted to a codeword in the double cover of the base normal graph.

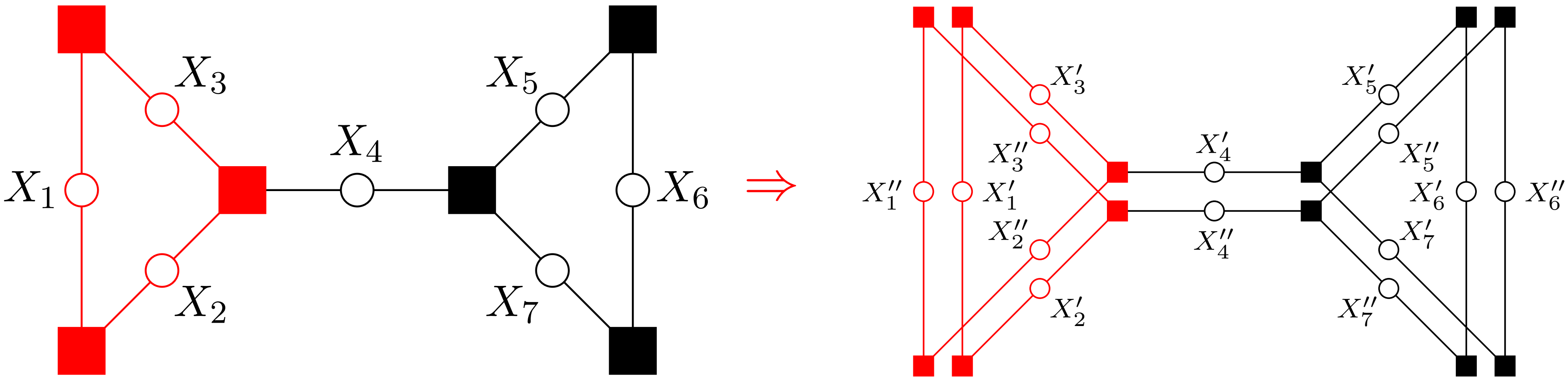


$\Rightarrow$

$$(1, 1, 1, 0, 0, 0, 0)$$

Codewords in Graph Covers

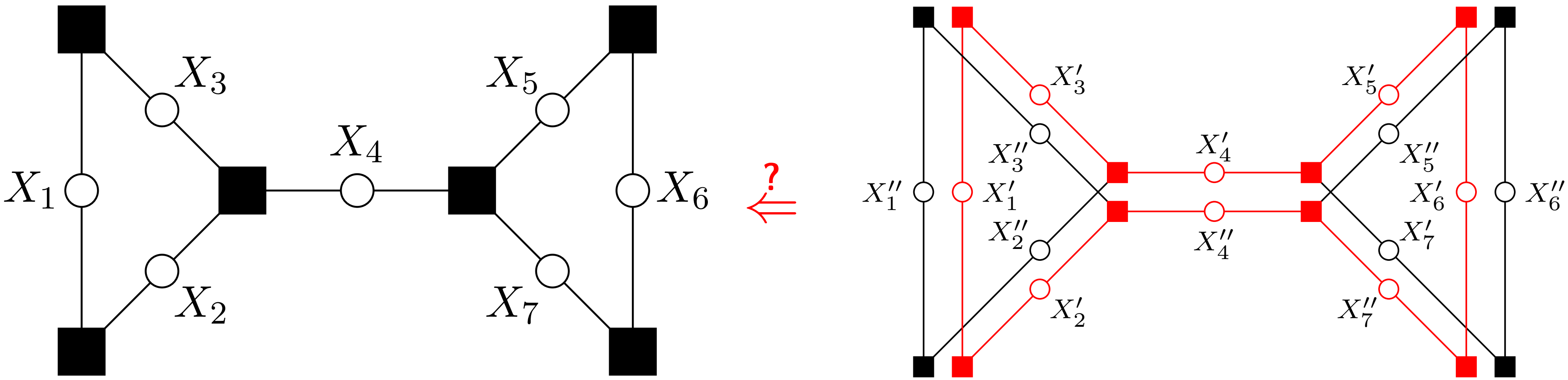
Obviously, any codeword in the base normal factor graph can be lifted to a codeword in the double cover of the base normal graph.



$(1, 1, 1, 0, 0, 0, 0)$ $\quad$ $(1{:}1,\ 1{:}1,\ 1{:}1,\ 0{:}0,\ 0{:}0,\ 0{:}0,\ 0{:}0)$

Codewords in Graph Covers

But in the double cover of the base normal factor graph there are also codewords that are not liftings of codewords in the base factor graph!

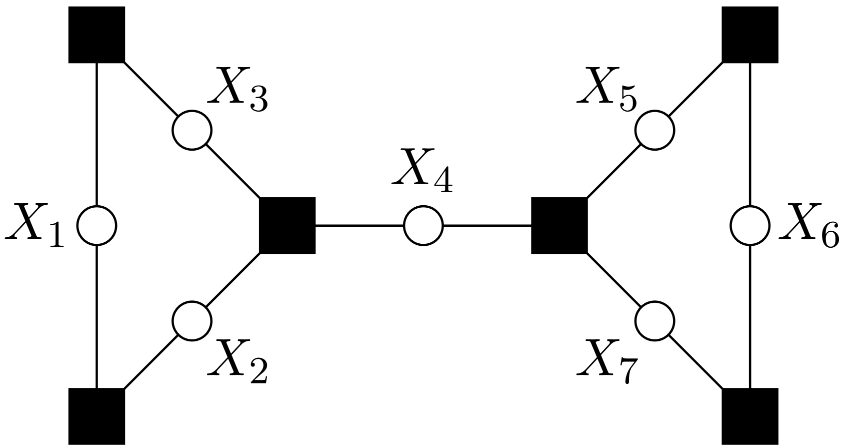


?

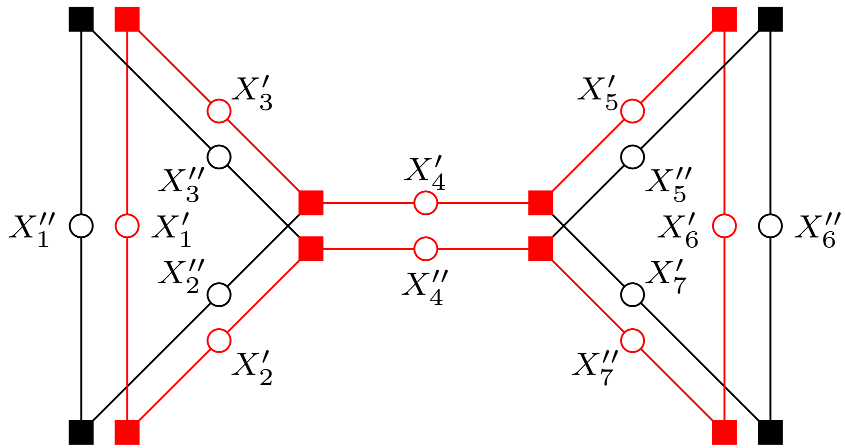
$(1{:}0,\ 1{:}0,\ 1{:}0,\ 1{:}1,\ 1{:}0,\ 1{:}0,\ 0{:}1)$

Codewords in Graph Covers

But in the double cover of the base normal factor graph there are also codewords that are not liftings of codewords in the base factor graph!



$\overset{?}{\Longleftarrow}$



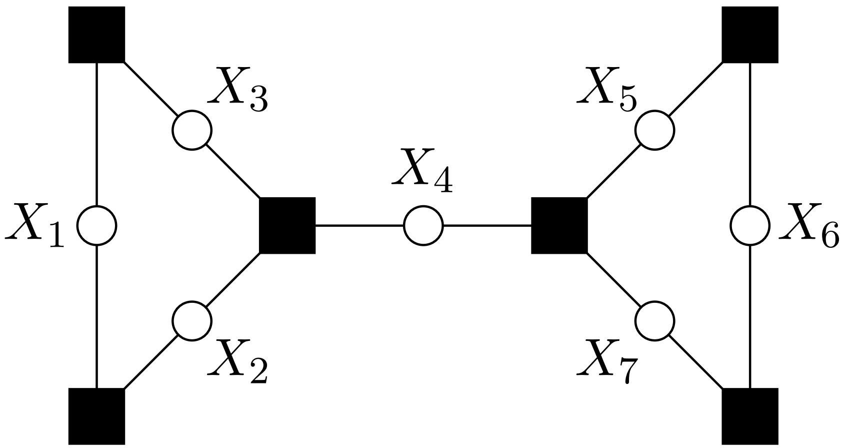
What about

$$\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{2}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)\ ?$$

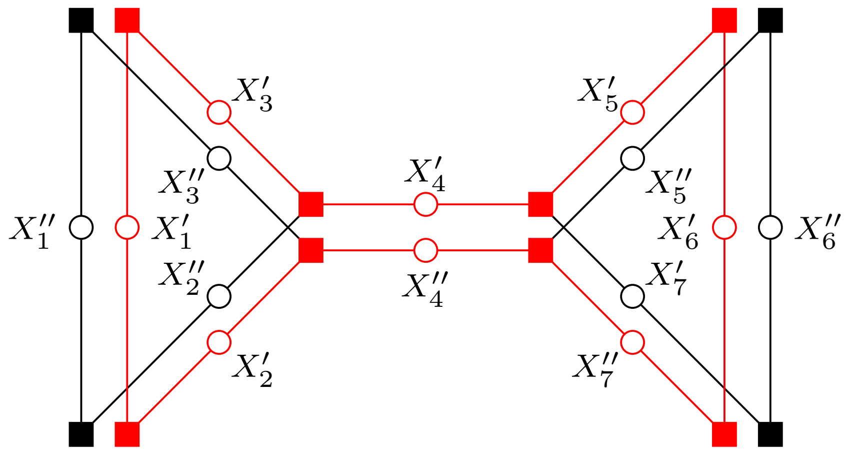
$$(1{:}0,\ 1{:}0,\ 1{:}0,\ 1{:}1,\ 1{:}0,\ 1{:}0,\ 0{:}1)$$

Codewords in Graph Covers

But in the double cover of the base normal factor graph there are also codewords that are not liftings of codewords in the base factor graph!



$\overset{?}{\Leftarrow}$



What about

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{2}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) ?$$

$$(1{:}0,\ 1{:}0,\ 1{:}0,\ 1{:}1,\ 1{:}0,\ 1{:}0,\ 0{:}1)$$

$\Rightarrow$ We will call such a vector a (graph-cover) pseudo-codeword.

Codewords in Graph Covers

Theorem:

Codewords in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

Codewords in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

- Let $\mathcal{P} \triangleq \mathcal{P}(\mathbf{H})$ be the first-order LP relaxation of $\mathrm{conv}(\mathcal{C})$ (aka fundamental polytope of $\mathbf{H}$).

Codewords in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

- Let $\mathcal{P} \triangleq \mathcal{P}(\mathbf{H})$ be the first-order LP relaxation of $\mathrm{conv}(\mathcal{C})$ (aka fundamental polytope of $\mathbf{H}$).
- Let $\mathcal{P}'$ be the set of all pseudo-codewords obtained through codewords in finite covers.

Codewords in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

- Let $\mathcal{P} \triangleq \mathcal{P}(\mathbf{H})$ be the first-order LP relaxation of $\mathrm{conv}(\mathcal{C})$ (aka fundamental polytope of $\mathbf{H}$).
- Let $\mathcal{P}'$ be the set of all pseudo-codewords obtained through codewords in finite covers.

Then $\mathcal{P}'$ is dense in $\mathcal{P}$, i.e.

$$\mathcal{P}' = \mathcal{P} \cap \mathbb{Q}^n.$$

Codewords in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

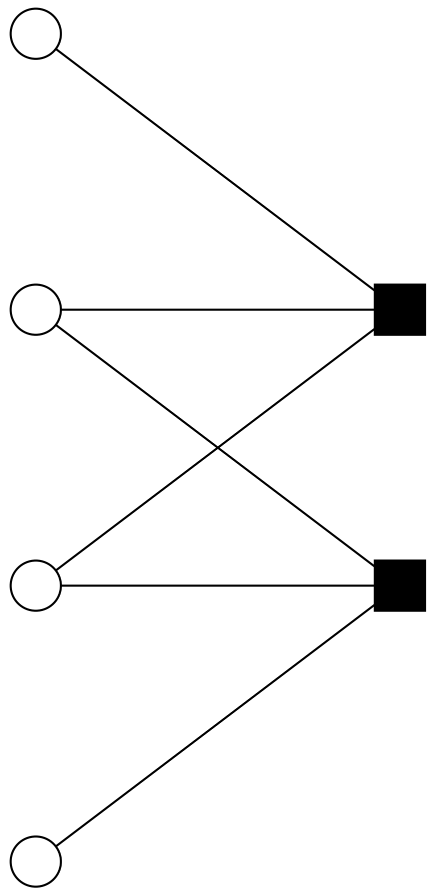
- Let $\mathcal{P} \triangleq \mathcal{P}(\mathbf{H})$ be the first-order LP relaxation of $\mathrm{conv}(\mathcal{C})$ (aka fundamental polytope of $\mathbf{H}$).
- Let $\mathcal{P}'$ be the set of all pseudo-codewords obtained through codewords in finite covers.

Then $\mathcal{P}'$ is dense in $\mathcal{P}$, i.e.

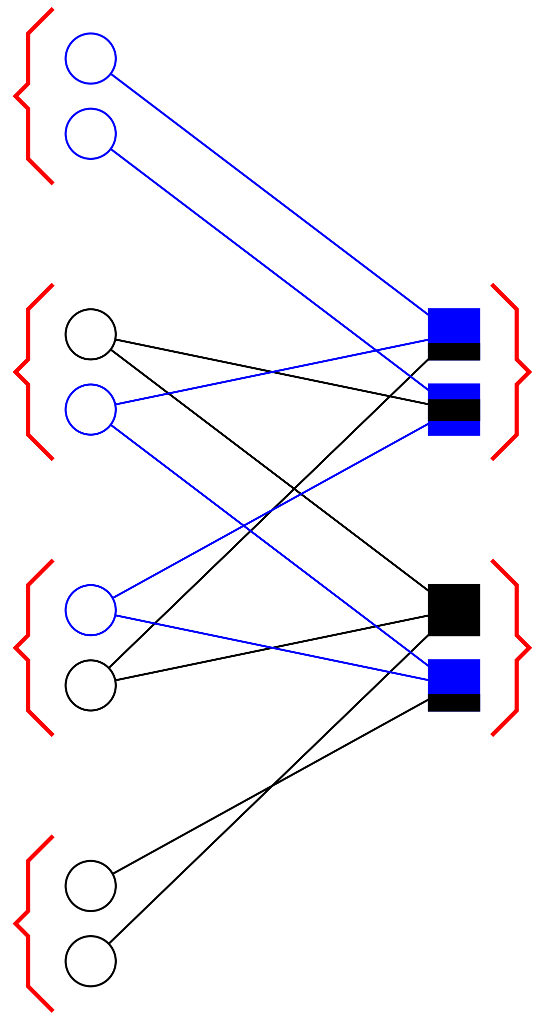
$$\mathcal{P}' = \mathcal{P} \cap \mathbb{Q}^n.$$

Moreover, note that all vertices of $\mathcal{P}$ are vectors with rational entries and are therefore also in $\mathcal{P}'$.

Valid Configurations in Graph Covers



Valid Configurations in Graph Covers



Valid Configurations in Graph Covers

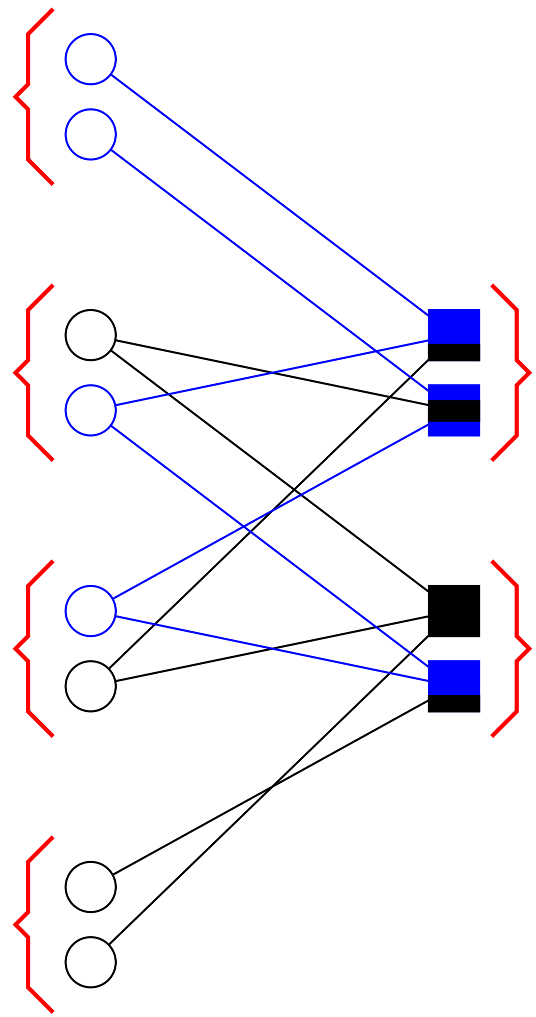
The components of the pseudo-marginal

$$\boldsymbol{\alpha} = \left\{ \{\boldsymbol{\alpha}_i\}_{i \in \mathcal{I}},\ \{\boldsymbol{\alpha}_j\}_{j \in \mathcal{J}} \right\}$$

associated to the valid configuration $\tilde{\mathbf{x}}$ are defined to be

$$\boldsymbol{\alpha}_i \triangleq \frac{1}{M} \sum_{m \in [M]} \tilde{\mathbf{x}}_{i,m},$$

$$\boldsymbol{\alpha}_j \triangleq \frac{1}{M} \sum_{m \in [M]} \tilde{\mathbf{x}}_{j,m}.$$



Valid Configurations in Graph Covers

The components of the pseudo-marginal

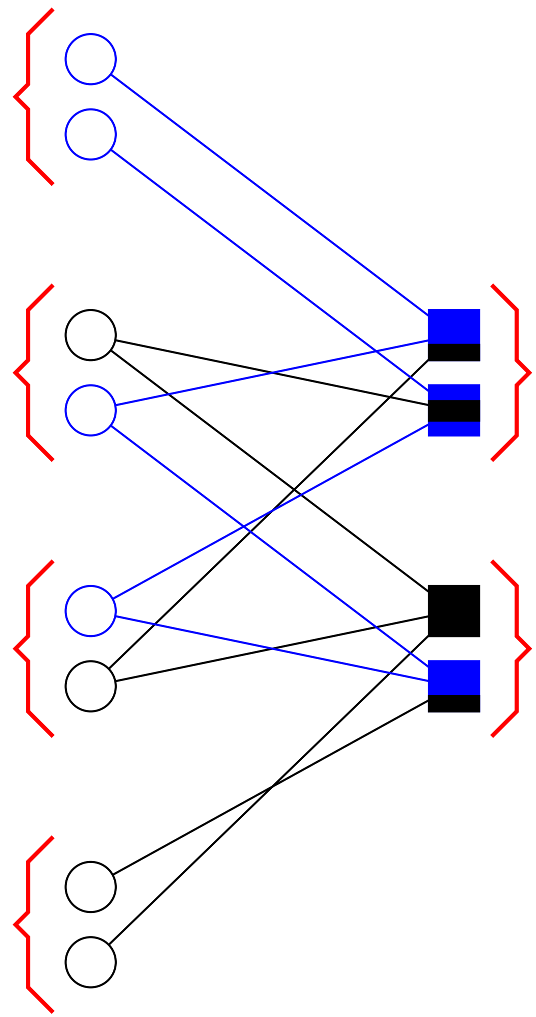
$$\boldsymbol{\alpha} = \Big\{ \{\boldsymbol{\alpha}_i\}_{i \in \mathcal{I}},\ \{\boldsymbol{\alpha}_j\}_{j \in \mathcal{J}} \Big\}$$

associated to the valid configuration $\tilde{\mathbf{x}}$ are defined to be

$$\boldsymbol{\alpha}_i \triangleq \frac{1}{M} \sum_{m \in [M]} \tilde{\mathbf{x}}_{i,m},$$

$$\boldsymbol{\alpha}_j \triangleq \frac{1}{M} \sum_{m \in [M]} \tilde{\mathbf{x}}_{j,m}.$$

The mapping from any M-fold graph cover to the base graph will be called $\boldsymbol{\varphi}_M$.



Valid Configurations in Graph Covers

Theorem:

Valid Configurations in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

Valid Configurations in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

- Let $\mathcal{L} \triangleq \mathcal{L}(\mathbf{H})$ be the local marginal poltyope of the factor graph corresponding to $\mathbf{H}$.

Valid Configurations in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

- Let $\mathcal{L} \triangleq \mathcal{L}(\mathbf{H})$ be the local marginal poltyope of the factor graph corresponding to $\mathbf{H}$.
- Let $\mathcal{L}'$ be the set of all pseudo-marginals obtained through codewords in finite covers.

Valid Configurations in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

- Let $\mathcal{L} \triangleq \mathcal{L}(\mathbf{H})$ be the local marginal poltyope of the factor graph corresponding to $\mathbf{H}$.
- Let $\mathcal{L}'$ be the set of all pseudo-marginals obtained through codewords in finite covers.

Then $\mathcal{L}'$ is dense in $\mathcal{L}$, i.e.

$$\mathcal{L}' = \mathcal{L} \cap \mathbb{Q}^{\dim(\mathcal{L})}.$$

Valid Configurations in Graph Covers

Theorem:

Consider a binary linear $\mathcal{C}$ defined by the parity-check matrix $\mathbf{H}$.

- Let $\mathcal{L} \triangleq \mathcal{L}(\mathbf{H})$ be the local marginal poltyope of the factor graph corresponding to $\mathbf{H}$.
- Let $\mathcal{L}'$ be the set of all pseudo-marginals obtained through codewords in finite covers.

Then $\mathcal{L}'$ is dense in $\mathcal{L}$, i.e.

$$\mathcal{L}' = \mathcal{L} \cap \mathbb{Q}^{\dim(\mathcal{L})}.$$

Moreover, note that all vertices of $\mathcal{L}$ are vectors with rational entries and are therefore also in $\mathcal{L}'$.

The Mapping φ_M

The Mapping φ_M

More formally, for any positive integer M, φ_M is the mapping

$$\begin{aligned}\varphi_M : \left\{ (\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}) \;\middle|\; \widetilde{\mathsf{G}} \in \widetilde{\mathcal{G}}_M,\ \widetilde{\mathbf{x}} \text{ is a valid configuration in } \widetilde{\mathcal{G}}_M \right\} &\longrightarrow \mathcal{L} \\ (\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}) &\longmapsto \boldsymbol{\alpha}\end{aligned}$$

The Mapping $\boldsymbol{\varphi}_M$

More formally, for any positive integer M, $\boldsymbol{\varphi}_M$ is the mapping

$$\begin{aligned} \boldsymbol{\varphi}_M : \Big\{ \big(\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}\big) \;\Big|\; \widetilde{\mathsf{G}} \in \widetilde{\mathcal{G}}_M,\; \widetilde{\mathbf{x}} \text{ is a valid configuration in } \widetilde{\mathcal{G}}_M \Big\} &\to \boldsymbol{\mathcal{L}} \\ \big(\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}\big) &\mapsto \boldsymbol{\alpha} \end{aligned}$$

where $\widetilde{\mathcal{G}}_M$ is the set of all M-fold covers of the base graph G.

The Mapping $\boldsymbol{\varphi}_M$

More formally, for any positive integer M, $\boldsymbol{\varphi}_M$ is the mapping

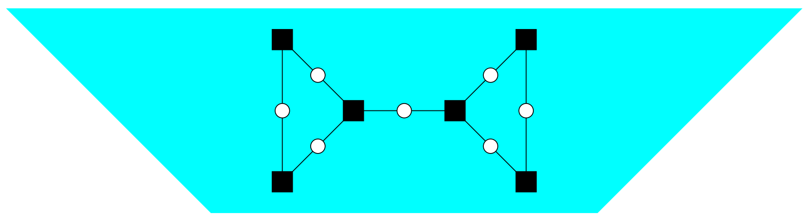
$$\begin{aligned} \boldsymbol{\varphi}_M : \Big\{ \big(\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}\big) \;\Big|\; \widetilde{\mathsf{G}} \in \widetilde{\mathcal{G}}_M,\ \widetilde{\mathbf{x}} \text{ is a valid configuration in } \widetilde{\mathcal{G}}_M \Big\} &\to \boldsymbol{\mathcal{L}} \\ \big(\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}\big) &\mapsto \boldsymbol{\alpha} \end{aligned}$$

where $\widetilde{\mathcal{G}}_M$ is the set of all M-fold covers of the base graph G.

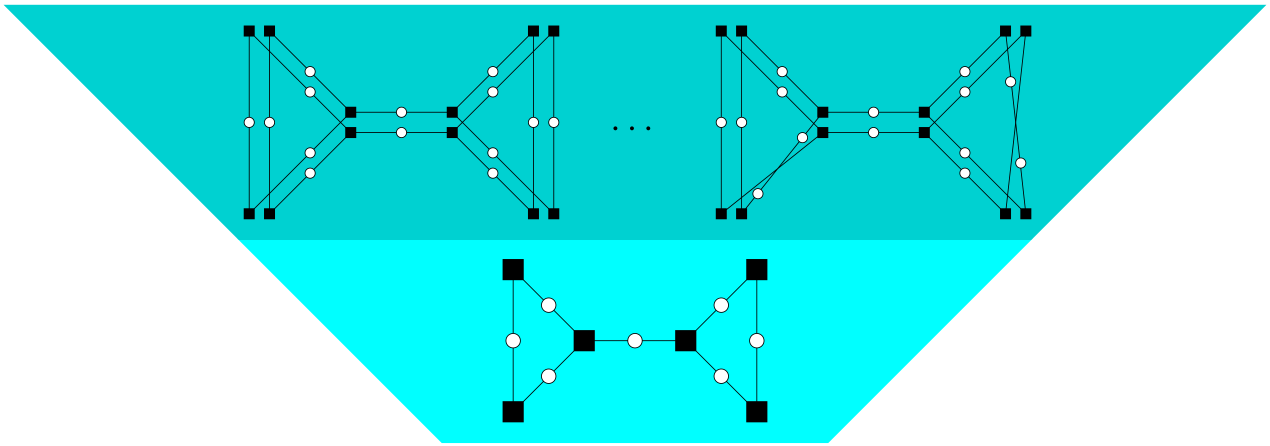
Important: the graph $\widetilde{\mathsf{G}} \in \widetilde{\mathcal{G}}_M$ needs to be included in the tuple $\big(\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}\big)$, otherwise $\widetilde{\mathbf{x}}$ is not well defined. This is especially crucial once we consider the inverse mapping $\boldsymbol{\varphi}_M^{-1}$.

Counting codewords in graph covers

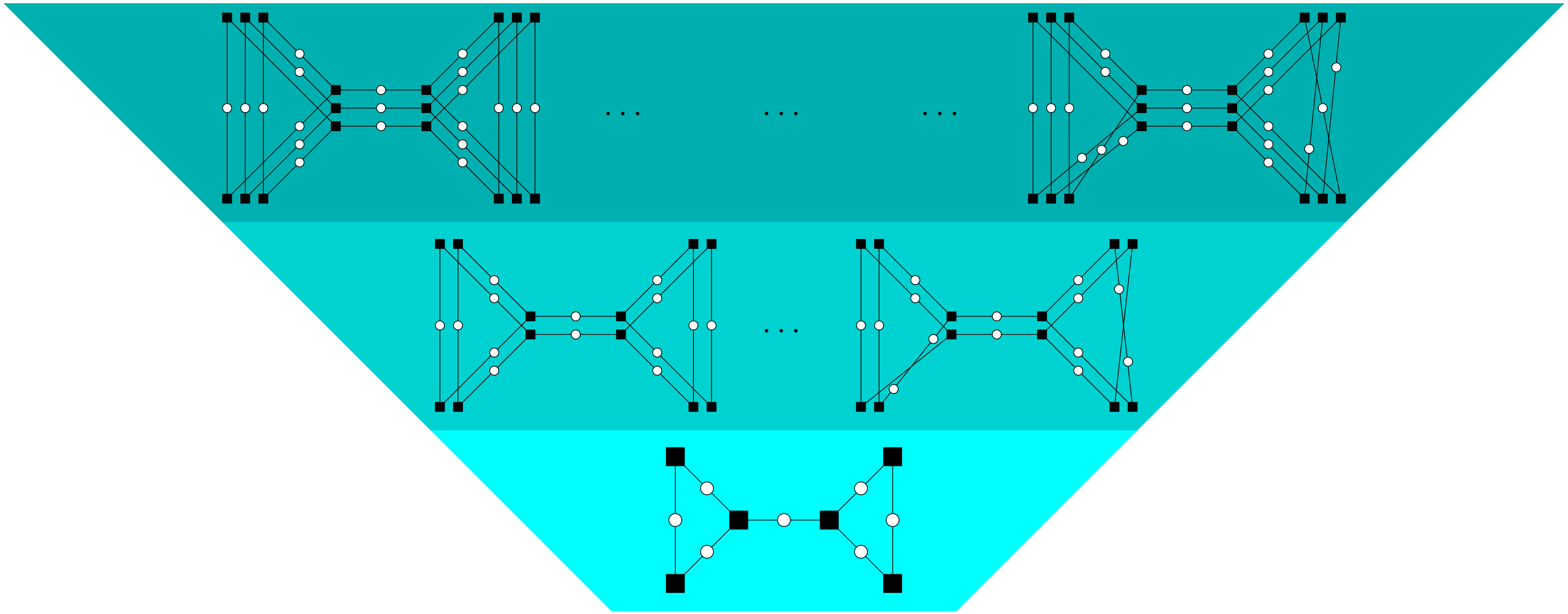
Graph Cover Hierarchy



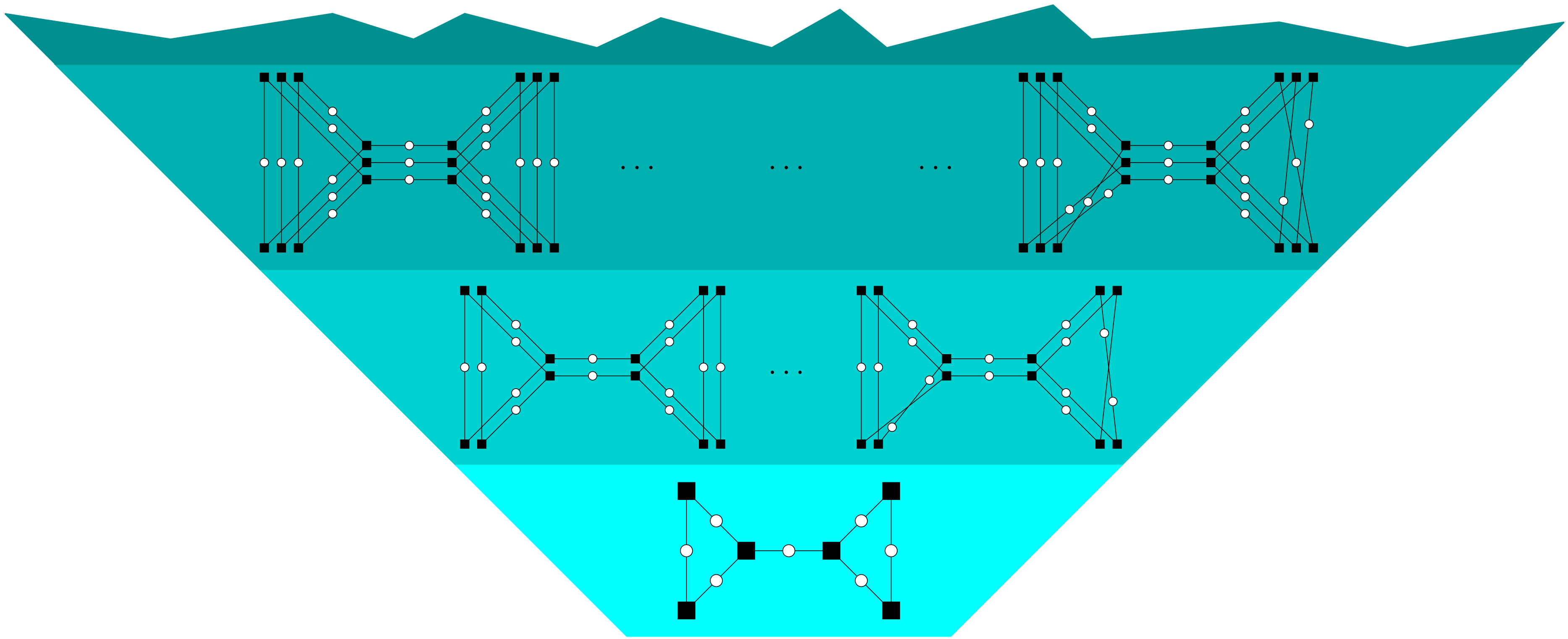
Graph Cover Hierarchy



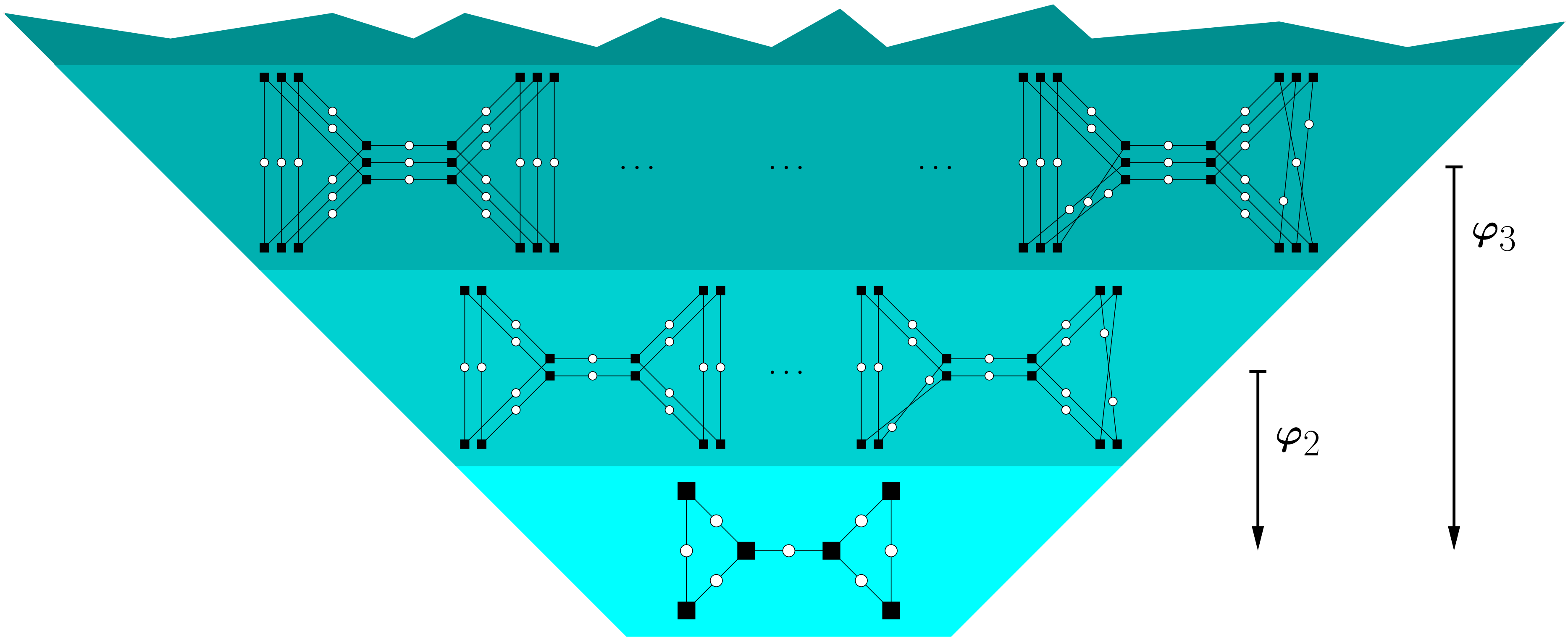
Graph Cover Hierarchy



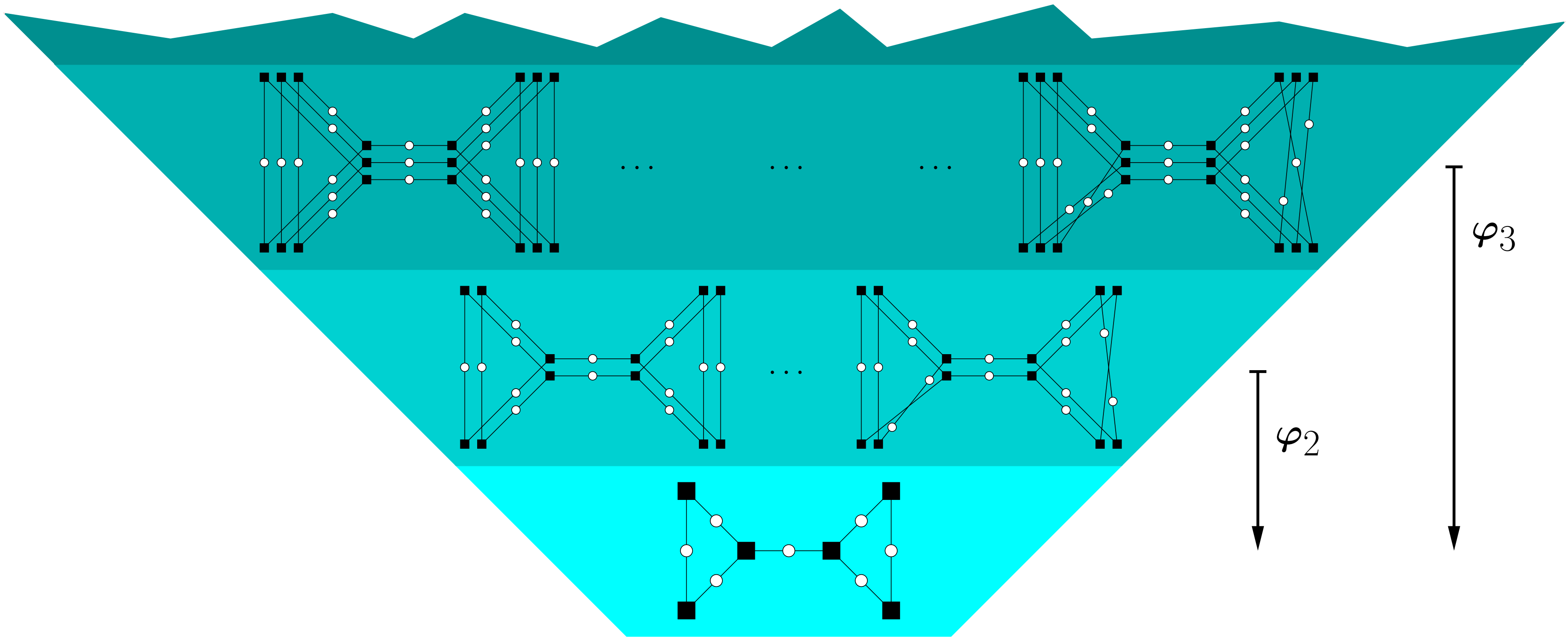
Graph Cover Hierarchy



Graph Cover Hierarchy

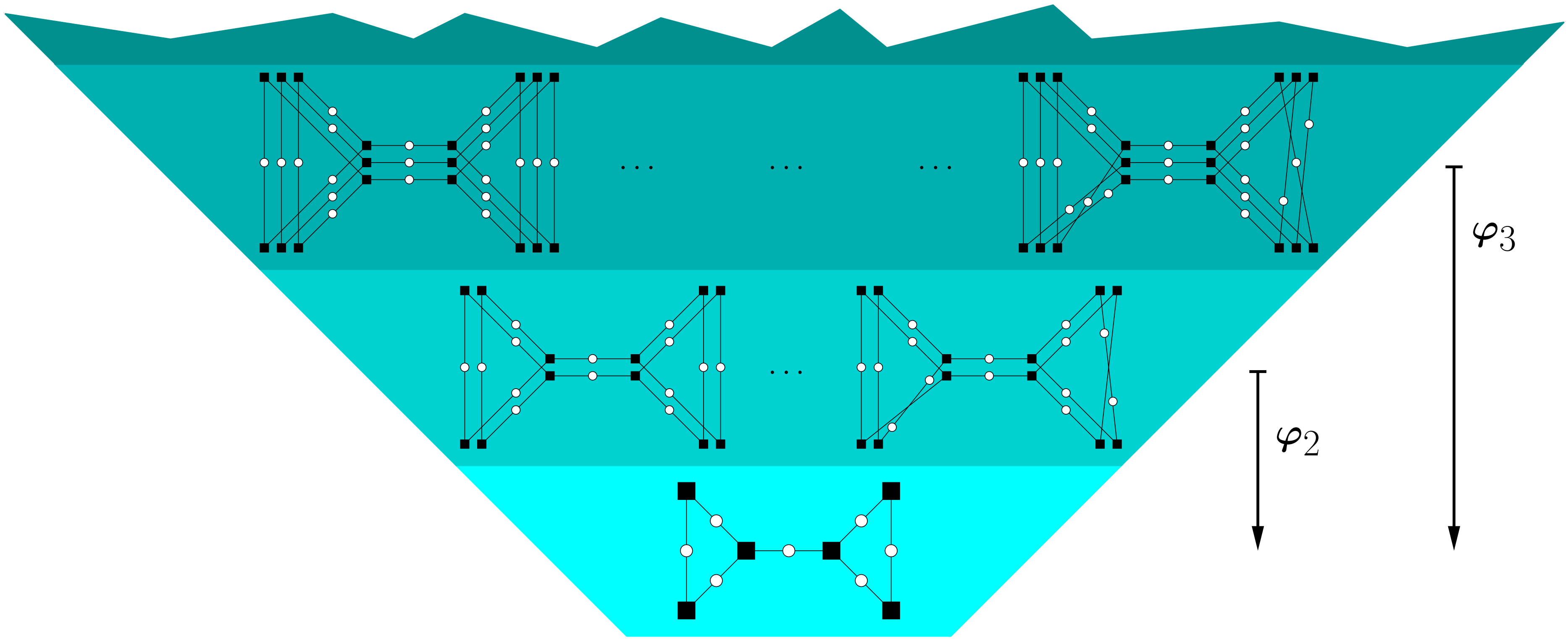


Graph Cover Hierarchy



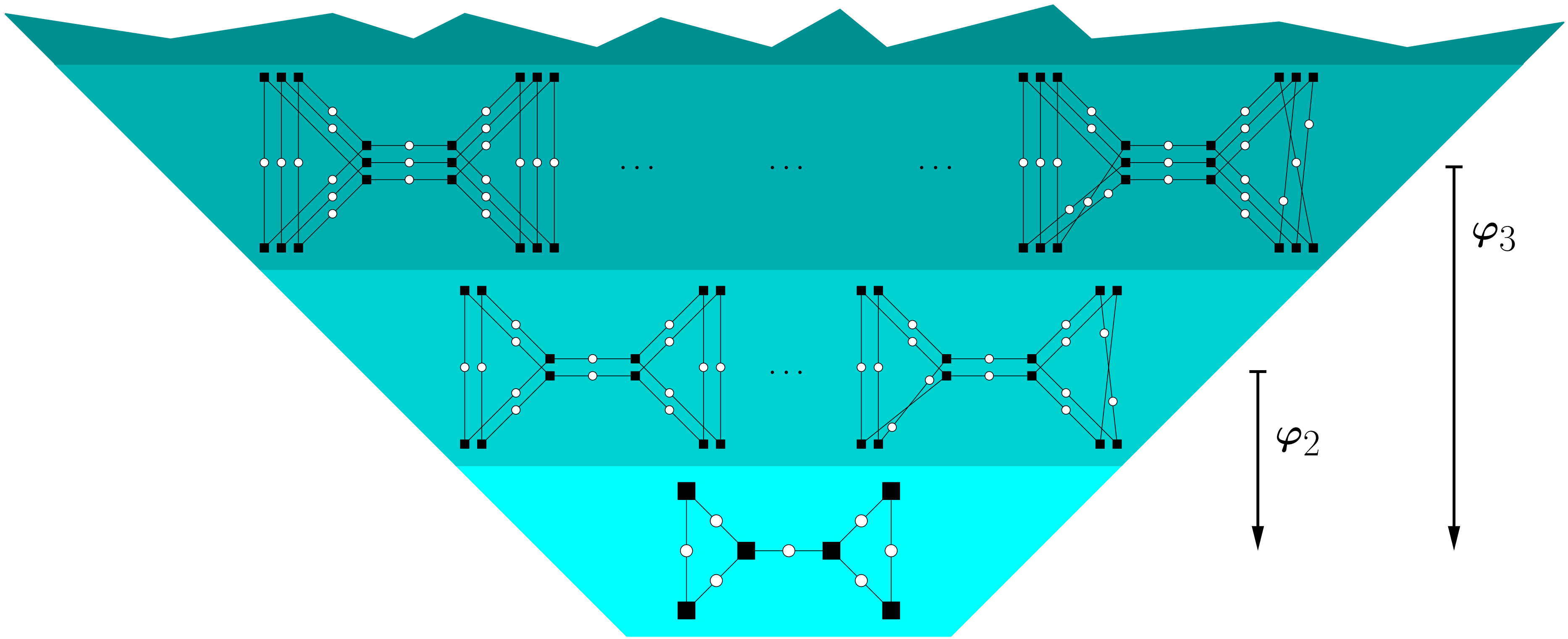
Let $\alpha \in \mathcal{L}'$ be a pseudo-marginal.

Graph Cover Hierarchy



Let $\alpha \in \mathcal{L}'$ be a pseudo-marginal. Then consider:

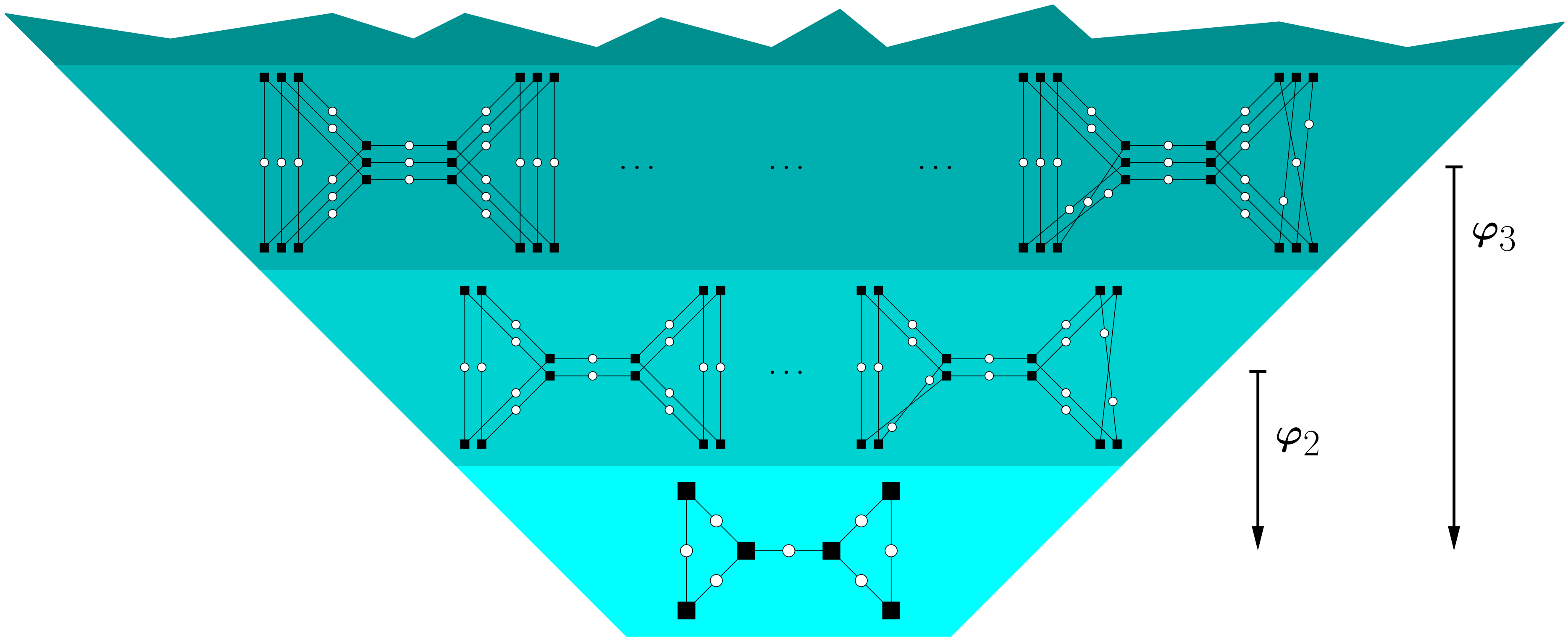
Graph Cover Hierarchy



Let $\boldsymbol{\alpha} \in \mathcal{L}'$ be a pseudo-marginal. Then consider:

$$\#\boldsymbol{\varphi}_M^{-1}(\boldsymbol{\alpha})$$

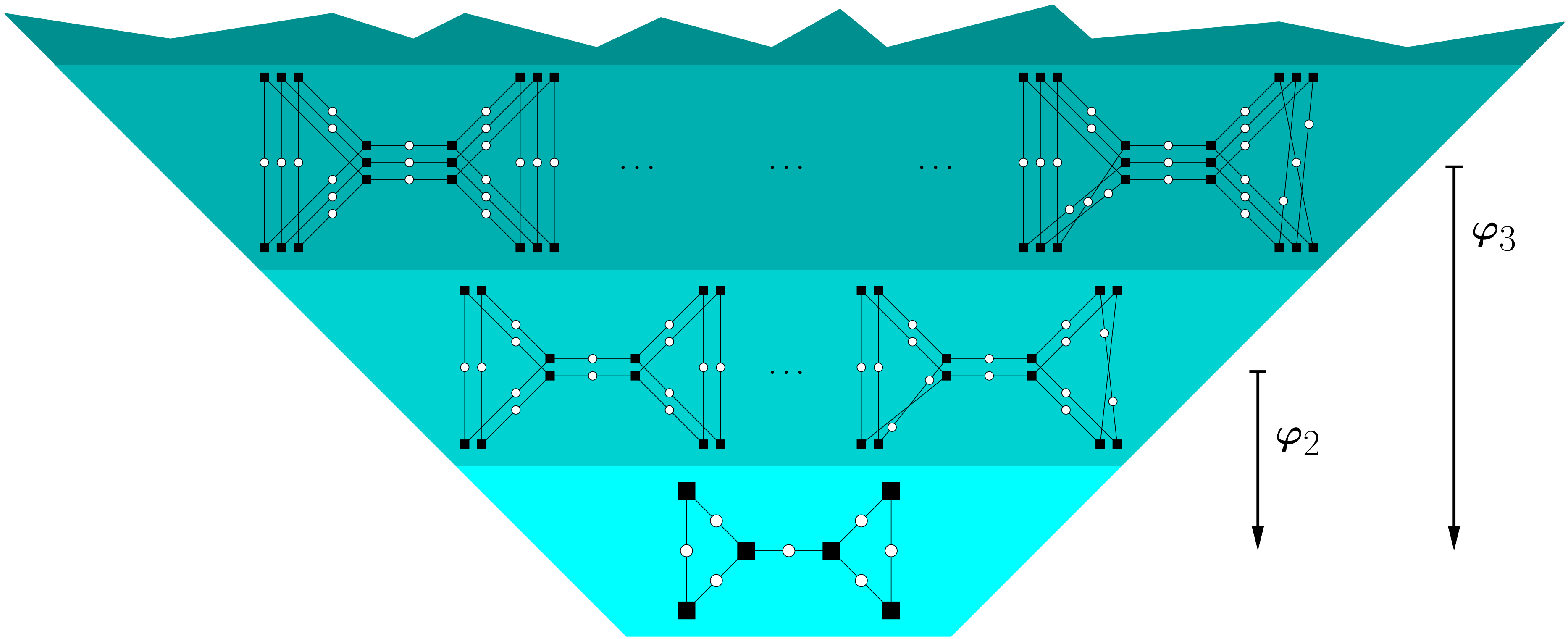
Graph Cover Hierarchy



Let $\boldsymbol{\alpha} \in \mathcal{L}'$ be a pseudo-marginal. Then consider:

$$\frac{\#\boldsymbol{\varphi}_M^{-1}(\boldsymbol{\alpha})}{\#\widetilde{\mathcal{G}}_M}$$

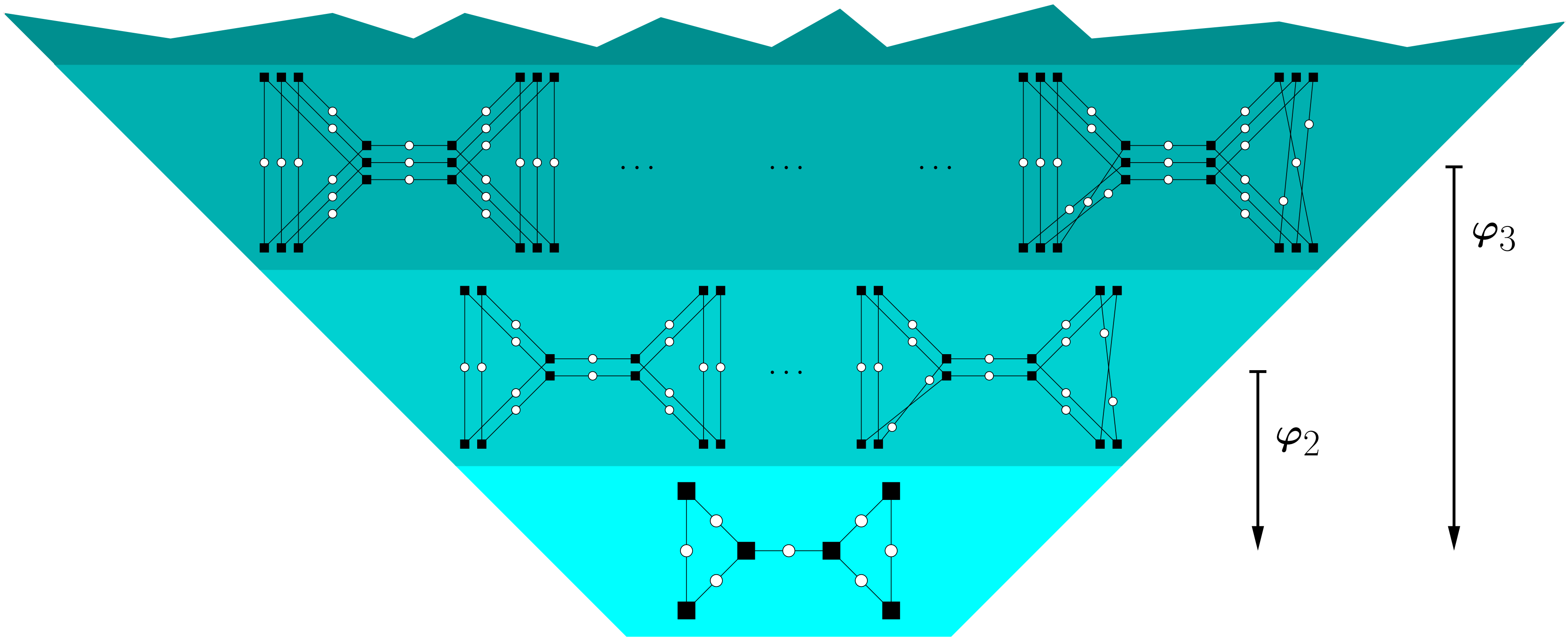
Graph Cover Hierarchy



Let $\boldsymbol{\alpha} \in \mathcal{L}'$ be a pseudo-marginal. Then consider:

$$\frac{1}{M} \log \frac{\#\boldsymbol{\varphi}_M^{-1}(\boldsymbol{\alpha})}{\#\widetilde{\mathcal{G}}_M}$$

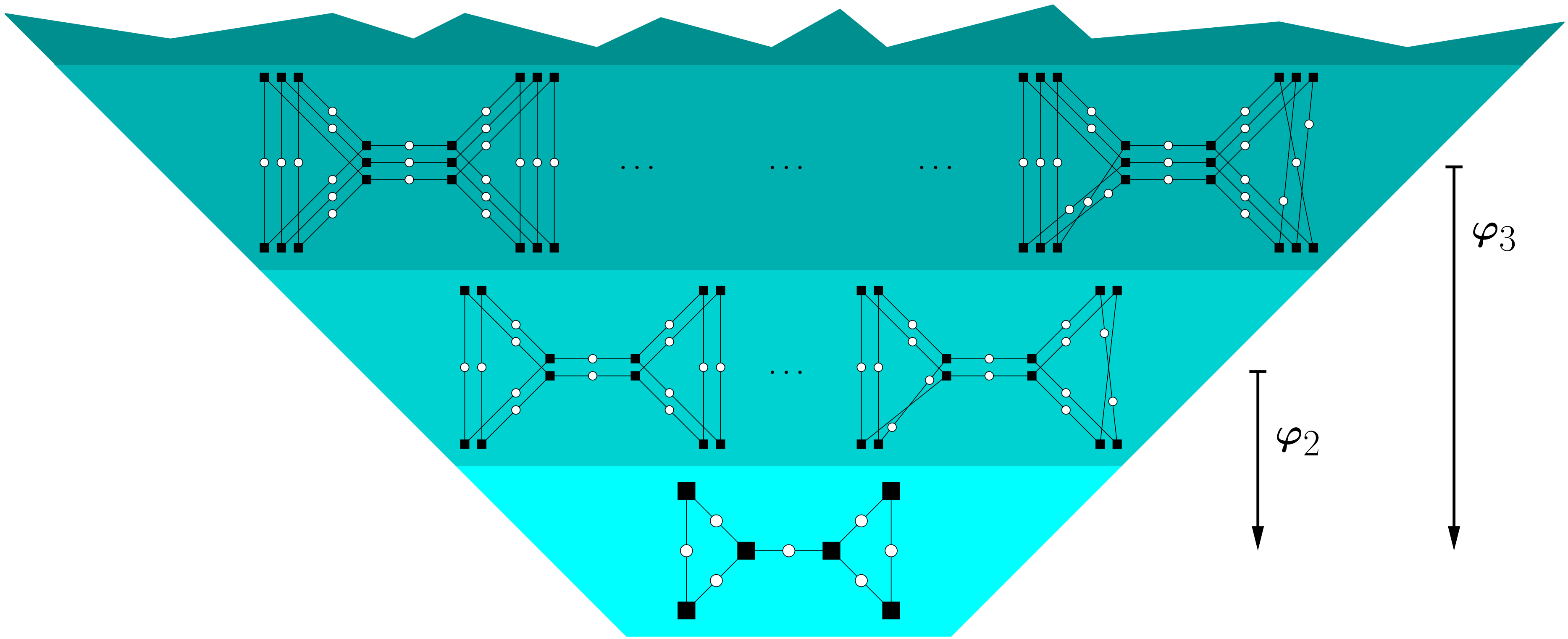
Graph Cover Hierarchy



Let $\boldsymbol{\alpha} \in \mathcal{L}'$ be a pseudo-marginal. Then consider:

$$\limsup_{M \to \infty} \frac{1}{M} \log \frac{\# \boldsymbol{\varphi}_M^{-1}(\boldsymbol{\alpha})}{\# \widetilde{\mathcal{G}}_M}$$

Graph Cover Hierarchy



Let $\boldsymbol{\alpha} \in \mathcal{L}'$ be a pseudo-marginal.

$$\textbf{Theorem:} \quad \limsup_{M \to \infty} \frac{1}{M} \log \frac{\# \boldsymbol{\varphi}_M^{-1}(\boldsymbol{\alpha})}{\# \widetilde{\mathcal{G}}_M} = H_{\mathrm{Bethe}}(\boldsymbol{\alpha})$$

Comments on the Previous Theorem

Comments on the Previous Theorem

- Similarly to the computation of the asymptotic growth rate of average Hamming spectra one has to be somewhat careful in formulating the above limit; we leave out the details.

Comments on the Previous Theorem

- Similarly to the computation of the asymptotic growth rate of average Hamming spectra one has to be somewhat careful in formulating the above limit; we leave out the details.

- Note: The ratio

$$\frac{\#\boldsymbol{\varphi}_M^{-1}(\boldsymbol{\alpha})}{\#\widetilde{\mathcal{G}}_M}$$

  represents the average number of valid configurations $\tilde{\mathbf{x}}$ per M-fold cover with associated pseudo-marginal $\boldsymbol{\alpha}$. Therefore, $H_{\mathrm{Bethe}}(\boldsymbol{\alpha})$ gives the asymptotic growth rate of that quantity.

Comments on the Previous Theorem

Comments on the Previous Theorem

- The above result is based on similar computations as in the derivation of the asymptotic growth rate of the average Hamming weight of protograph-based LDPC codes. Cf.
  - [Fogal/McEliece/Thorpe, 2005],
  - papers by Divsalar, Ryan, et al. (2005–).

Comments on the Previous Theorem

- The above result is based on similar computations as in the derivation of the asymptotic growth rate of the average Hamming weight of protograph-based LDPC codes. Cf.
  - [Fogal/McEliece/Thorpe, 2005],
  - papers by Divsalar, Ryan, et al. (2005–).

- To the best of our knowledge, the above interpretation of the Bethe entropy cannot be found in the literature (besides the talks that we gave at the 2008 Allerton Conference / 2009 ITA Workshop in San Diego).

A “micro-/macrostate reinterpretation” of the theorem by Yedidia et al.

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

With the definitions so far, let us consider the following setup.

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

With the definitions so far, let us consider the following setup.

- Fix some positive integer M.
  (Finally, we are mostly interested in the limit $M \to \infty$.)

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

With the definitions so far, let us consider the following setup.

- Fix some positive integer M.
  (Finally, we are mostly interested in the limit $M \to \infty$.)

- Set of microstates $\triangleq$ set of microstates$_M$

$$\triangleq \left((\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}) \;\middle|\; \widetilde{\mathsf{G}} \in \widetilde{\mathcal{G}}_M, \ \widetilde{\mathbf{x}} \text{ is a valid configuration in } \widetilde{\mathcal{G}}_M \right)$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

With the definitions so far, let us consider the following setup.

- Fix some positive integer M.
  (Finally, we are mostly interested in the limit $M \to \infty$.)

- Set of $\mathrm{microstates} \triangleq$ set of $\mathrm{microstates}_M$

$$\triangleq \left((\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}) \;\middle|\; \widetilde{\mathsf{G}} \in \widetilde{\mathcal{G}}_M, \; \widetilde{\mathbf{x}} \text{ is a valid configuration in } \widetilde{\mathcal{G}}_M \right)$$

- Set of $\mathrm{macrostates} \triangleq$ set of $\mathrm{macrostates}_M$

$$\triangleq \boldsymbol{\varphi}_M(\text{set of } \mathrm{microstates})$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\mathrm{microstate}) = \mathsf{const} \qquad \text{for all } \mathrm{microstates}.$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) = \text{const} \qquad \text{for all microstates.}$$

Then

$$P(\text{macrostate}) \;\propto\; \#\{\text{microstate} : \; \boldsymbol{\varphi}_M(\text{microstate}) = \text{macrostate}\}$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) = \text{const} \qquad \text{for all microstates.}$$

Then

$$\begin{aligned} P(\text{macrostate}) &\propto \#\{\text{microstate} : \boldsymbol{\varphi}_M(\text{microstate}) = \text{macrostate}\} \\ &= \#\boldsymbol{\varphi}_M^{-1}(\text{macrostate}) \end{aligned}$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) = \mathsf{const} \qquad \text{for all microstates.}$$

Then

$$\begin{aligned} P(\text{macrostate}) &\propto \#\big\{\text{microstate} : \ \boldsymbol{\varphi}_M(\text{microstate}) = \text{macrostate}\big\} \\ &= \#\boldsymbol{\varphi}_M^{-1}(\text{macrostate}) \\ &\propto \frac{\#\boldsymbol{\varphi}_M^{-1}(\text{macrostate})}{\#\widetilde{\mathcal{G}}_M} \end{aligned}$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) = \text{const} \qquad \text{for all microstates.}$$

Then

$$\begin{aligned}
P(\text{macrostate}) &\propto \#\Big\{\text{microstate} : \ \boldsymbol{\varphi}_M(\text{microstate}) = \text{macrostate}\Big\} \\
&= \#\boldsymbol{\varphi}_M^{-1}(\text{macrostate}) \\
&\propto \frac{\#\boldsymbol{\varphi}_M^{-1}(\text{macrostate})}{\#\widetilde{\mathcal{G}}_M} \\
&\overset{\approx}{\propto} \exp\Big(M \cdot H_{\text{Bethe}}(\text{macrostate})\Big).
\end{aligned}$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) \;\propto\; \exp\Big(-M\cdot E\big(\boldsymbol{\varphi}_M(\text{microstate})\big)\Big) \quad \text{for all microstates.}$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) \;\propto\; \exp\Big(-M \cdot E\big(\boldsymbol{\varphi}_M(\text{microstate})\big)\Big) \quad \text{for all microstates.}$$

Then

$$P(\text{macrostate}) \;\propto\; \exp\big(-M \cdot E(\text{macrostate})\big) \cdot \#\{\text{microstate} : \; \varphi(\text{microstate}) = \text{macrostate}\}$$

"Micro-/Macrostate Reinterpretation" of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) \;\propto\; \exp\Big(-M\cdot E\big(\boldsymbol{\varphi}_M(\text{microstate})\big)\Big) \quad \text{for all microstates.}$$

Then

$$\begin{aligned} P(\text{macrostate}) \;&\propto\; \exp\big(-M\cdot E(\text{macrostate})\big) \\ &\qquad \cdot\ \#\big\{\text{microstate}:\ \varphi(\text{microstate})=\text{macrostate}\big\} \\ &=\; \exp\big(-M\cdot E(\text{macrostate})\big)\ \cdot\ \#\varphi^{-1}(\text{macrostate}) \end{aligned}$$

“Micro-/Macrostate Reinterpretation” of the Theorem by Yedidia et al.

Assume

$$P(\text{microstate}) \;\propto\; \exp\Big(-M \cdot E\big(\boldsymbol{\varphi}_M(\text{microstate})\big)\Big) \quad \text{for all microstates.}$$

Then

$$\begin{aligned}
P(\text{macrostate}) \;&\propto\; \exp\big(-M \cdot E(\text{macrostate})\big) \\
&\qquad \cdot \#\big\{\text{microstate} : \;\varphi(\text{microstate}) = \text{macrostate}\big\} \\
&= \exp\big(-M \cdot E(\text{macrostate})\big) \cdot \#\varphi^{-1}(\text{macrostate}) \\
&\mathrel{\overset{\approx}{\propto}} \exp\big(-M \cdot E(\text{macrost.})\big) \cdot \exp\big(M \cdot H_{\text{Bethe}}(\text{macrost.})\big)
\end{aligned}$$

Fixed Points of the SPA

Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the SPA correspond to stationary points of the Variational Bethe free energy (VBFE).

Fixed Points of the SPA

Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the SPA correspond to stationary points of the Variational Bethe free energy (VBFE).

Re-interpretation in terms of graph covers:

Let

$$P(\text{microstate}) \;\propto\; \exp\Big(-M \cdot \langle \boldsymbol{\varphi}_M(\text{microstate}), \boldsymbol{\lambda} \rangle\Big)$$

Fixed Points of the SPA

Theorem (Yedidia/Freeman/Weiss, 2000)
Fixed points of the SPA correspond to stationary points of the Variational Bethe free energy (VBFE).

Re-interpretation in terms of graph covers:

Let

$$P(\text{microstate}) \;\propto\; \exp\Big(-M \cdot \langle \boldsymbol{\varphi}_M(\text{microstate}), \boldsymbol{\lambda} \rangle\Big)$$

Then

$$P(\text{macrostate}) \;=\; \exp\Big(-M \cdot \langle \text{macrostate}, \boldsymbol{\lambda} \rangle\Big) \cdot \#\varphi^{-1}(\text{macrostate})$$

Fixed Points of the SPA

Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the SPA correspond to *stationary points* of the Variational Bethe free energy (VBFE).

Fixed Points of the SPA

Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the SPA correspond to *stationary points* of the Variational Bethe free energy (VBFE).

Re-interpretation in terms of graph covers:

A fixed point of the SPA corresponds to a macrostate $\boldsymbol{\alpha}$, i.e., a pseudo-marginal $\boldsymbol{\alpha}$, that is a *stationary point* of

$$P(\boldsymbol{\alpha}) \;\propto\; \exp\Big(-M \cdot \langle \boldsymbol{\alpha}, \boldsymbol{\lambda} \rangle\Big) \cdot \#\varphi^{-1}(\boldsymbol{\alpha})$$

when M goes to infinity.

Fixed Points of the SPA

Theorem (Yedidia/Freeman/Weiss, 2000)

Fixed points of the SPA correspond to *local minima* of the Variational Bethe free energy (VBFE).

Re-interpretation in terms of graph covers:

A fixed point of the SPA corresponds to a macrostate $\boldsymbol{\alpha}$, i.e., a pseudo-marginal $\boldsymbol{\alpha}$, that is a *local maximimum* of

$$P(\boldsymbol{\alpha}) \ \propto \ \exp\Big(-M \cdot \langle \boldsymbol{\alpha}, \boldsymbol{\lambda} \rangle\Big) \cdot \#\varphi^{-1}(\boldsymbol{\alpha})$$

when M goes to infinity.

The Transient Part of the SPA

Static vs. Dynamic Setup

Symbols: σ: microstate, Σ: macrostate.

Static vs. Dynamic Setup

Symbols: σ: microstate, Σ: macrostate.

Static setup:

Static vs. Dynamic Setup

Symbols: σ: microstate, Σ: macrostate.

Static setup:

$$P(\Sigma) \;\propto\; \exp\big(-M \cdot E(\Sigma)\big) \;\cdot\; \#\varphi^{-1}(\Sigma)$$

Static vs. Dynamic Setup

Symbols: σ: microstate, Σ: macrostate.

Static setup:

$$P(\Sigma) \;\propto\; \exp\big(-M \cdot E(\Sigma)\big) \;\cdot\; \#\varphi^{-1}(\Sigma)$$

Dynamic setup:

$$P\big(\Sigma(t+\Delta t) \,\big|\, \sigma(t)\big)$$
$$\propto\; \exp\Big(-M \cdot E\big(\Sigma(t+\Delta t) \,\big|\, \sigma(t)\big)\Big) \;\cdot\; \#\varphi^{-1}\big(\Sigma(t+\Delta t) \,\big|\, \sigma(t)\big)$$

Static vs. Dynamic Setup

Symbols: σ: microstate, Σ: macrostate.

Static setup:

$$P(\Sigma) \;\propto\; \exp\big(-M \cdot E(\Sigma)\big) \;\cdot\; \#\varphi^{-1}(\Sigma)$$

Dynamic setup:

$$P\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)$$
$$\propto\; \exp\Big(-M \cdot E\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)\Big) \;\cdot\; \#\varphi^{-1}\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)$$

"Better" dynamic setup:

$$P\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)$$
$$\propto\; \exp\Big(-M \cdot E\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)\Big) \;\cdot\; \#\varphi^{-1}\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)$$

Static vs. Dynamic Setup

Symbols: σ: microstate, Σ: macrostate.

Static setup: $\boxed{\text{models fix points of the SPA}}$

$$P(\Sigma) \;\propto\; \exp\big(-M \cdot E(\Sigma)\big) \;\cdot\; \#\varphi^{-1}(\Sigma)$$

Dynamic setup:

$$P\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)$$
$$\propto\; \exp\Big(-M \cdot E\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)\Big) \;\cdot\; \#\varphi^{-1}\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)$$

"Better" dynamic setup:

$$P\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)$$
$$\propto\; \exp\Big(-M \cdot E\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)\Big) \;\cdot\; \#\varphi^{-1}\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)$$

Static vs. Dynamic Setup

Symbols: σ: microstate, Σ: macrostate.

Static setup: $\boxed{\text{models fix points of the SPA}}$

$$P(\Sigma) \ \propto \ \exp\big(-M \cdot E(\Sigma)\big) \ \cdot \ \#\varphi^{-1}(\Sigma)$$

Dynamic setup:

$$P\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)$$
$$\propto \ \exp\Big(-M \cdot E\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)\Big) \ \cdot \ \#\varphi^{-1}\big(\Sigma(t+\Delta t) \mid \sigma(t)\big)$$

"Better" dynamic setup: $\boxed{\text{will model the transient part of the SPA}}$

$$P\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)$$
$$\propto \ \exp\Big(-M \cdot E\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)\Big) \ \cdot \ \#\varphi^{-1}\big(\Sigma(t+\Delta t) \mid \Sigma(t)\big)$$

Graph-Dynamical Systems

We want to show that the transient part of the SPA can be expressed in terms of a graph-dynamical system.

Graph-Dynamical Systems

We want to show that the transient part of the SPA can be expressed in terms of a graph-dynamical system.

Graph-dynamical system (e.g., [Prisner:95]):

Graph-Dynamical Systems

We want to show that the transient part of the SPA can be expressed in terms of a graph-dynamical system.

Graph-dynamical system (e.g., [Prisner:95]):

- Let Γ be a set of graphs.

Graph-Dynamical Systems

We want to show that the transient part of the SPA can be expressed in terms of a graph-dynamical system.

Graph-dynamical system (e.g., [Prisner:95]):

- Let Γ be a set of graphs.
- Let Ψ be some (possibly random) mapping from Γ to Γ.

Graph-Dynamical Systems

We want to show that the transient part of the SPA can be expressed in terms of a graph-dynamical system.

Graph-dynamical system (e.g., [Prisner:95]):

- Let Γ be a set of graphs.
- Let Ψ be some (possibly random) mapping from Γ to Γ.
- Because the domain and the range of Ψ are equal, it makes sense to study the repeated application of the mapping Ψ:

$$\Gamma \quad \xrightarrow{\Psi} \quad \Gamma \quad \xrightarrow{\Psi} \quad \cdots \quad \xrightarrow{\Psi} \quad \Gamma$$

Review

(of the setup used in the re-interpretation of f.p.s of the SPA)

- Set of microstates

$$\triangleq \left((\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}) \;\middle|\; \widetilde{\mathsf{G}} \in \widetilde{\mathcal{G}}_M, \; \widetilde{\mathbf{x}} \text{ is a valid configuration in } \widetilde{\mathcal{G}}_M \right)$$

- Mapping $\boldsymbol{\varphi}_M$

$$\text{maps} \quad (\widetilde{\mathsf{G}}, \widetilde{\mathbf{x}}) \quad \text{to} \quad \boldsymbol{\omega}(\widetilde{\mathbf{x}})$$

- Set of macrostates

$$\triangleq \boldsymbol{\varphi}_M(\text{set of microstates})$$

Corresponding Setup for the Transient Part of the SPA

- Set of microstates

???

- Mapping φ_M

???

- Set of macrostates

???

Corresponding Setup for the Transient Part of the SPA

- Set of microstates

$$???$$

- Mapping φ_M

$$???$$

- Set of macrostates

$$???$$

Note: $\Gamma =$ set of M-covers of G and valid configurations therein is obviously not sufficient.

Corresponding Setup for the Transient Part of the SPA

- Set of microstates

  $\Rightarrow \ \Gamma =$ set of what we call colored hypergraph M-cover

  or colored twisted M-cover

- Mapping φ_M

  ???

- Set of macrostates

  ???

Corresponding Setup for the Transient Part of the SPA

- Set of microstates

  $\Rightarrow$ $\Gamma =$ set of what we call colored hypergraph M-cover
  or colored twisted M-cover

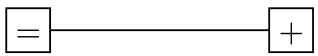
- Mapping φ_M

  ???

- Set of macrostates

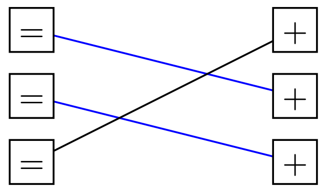
  set of all possible marginals on the LHS function nodes
  $\times$ set of all possible marginals on the RHS function nodes

Comment on Microstates



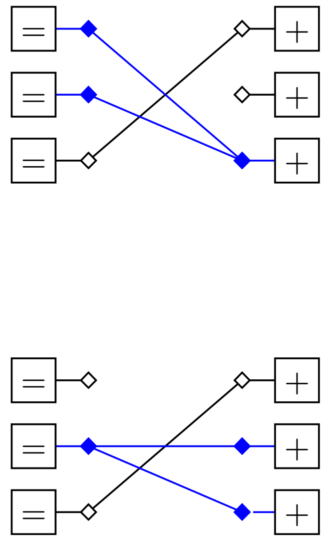
edge

in FFG



corrsponding edges

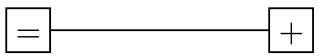
in some colored 3-cover



corresponding edges

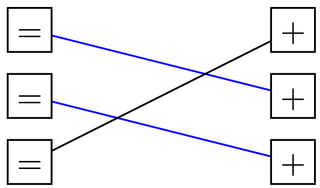
in colored hypergraph 3-cover

Comment on Microstates



edge

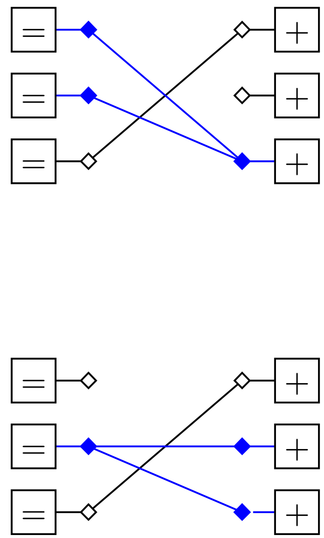
in FFG



corrsponding edges

in some colored 3-cover

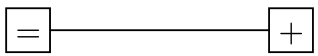
LHS and RHS marginals must match



corresponding edges

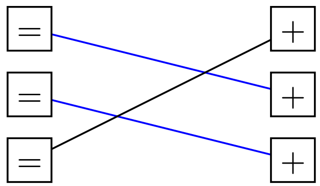
in colored hypergraph 3-cover

Comment on Microstates



edge

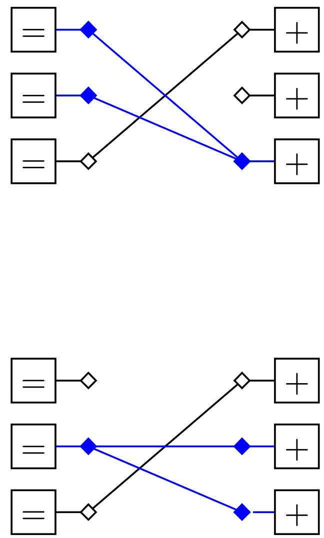
in FFG



corrsponding edges

in some colored 3-cover

LHS and RHS marginals
must match



corresponding edges

in colored hypergraph 3-cover

LHS and RHS marginals
do not have to match

Comment on Macrostates

$$
\begin{tikzcd}
\overleftarrow{\boldsymbol{\mu}}(0) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(0.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(1) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(1.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(2) \arrow[r] \arrow[d] & {} \\
\boldsymbol{\eta}_{\mathrm{left}}(0) & \boldsymbol{\eta}_{\mathrm{right}}(0.5) & \boldsymbol{\eta}_{\mathrm{left}}(1) & \boldsymbol{\eta}_{\mathrm{right}}(1.5) & \boldsymbol{\eta}_{\mathrm{left}}(2) & .
\end{tikzcd}
$$

Comment on Macrostates

$$
\begin{tikzcd}
\overleftarrow{\boldsymbol{\mu}}(0) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(0.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(1) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(1.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(2) \arrow[r] \arrow[d] & {} \\
\boldsymbol{\eta}_{\mathrm{left}}(0) & \boldsymbol{\eta}_{\mathrm{right}}(0.5) & \boldsymbol{\eta}_{\mathrm{left}}(1) & \boldsymbol{\eta}_{\mathrm{right}}(1.5) & \boldsymbol{\eta}_{\mathrm{left}}(2) &
\end{tikzcd}
$$

Comment on Macrostates

$$
\begin{tikzcd}
\overleftarrow{\boldsymbol{\mu}}(0) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(0.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(1) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(1.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(2) \arrow[r] \arrow[d] & {} \\
\boldsymbol{\eta}_{\text{left}}(0) & \boldsymbol{\eta}_{\text{right}}(0.5) & \boldsymbol{\eta}_{\text{left}}(1) & \boldsymbol{\eta}_{\text{right}}(1.5) & \boldsymbol{\eta}_{\text{left}}(2) &
\end{tikzcd}
$$

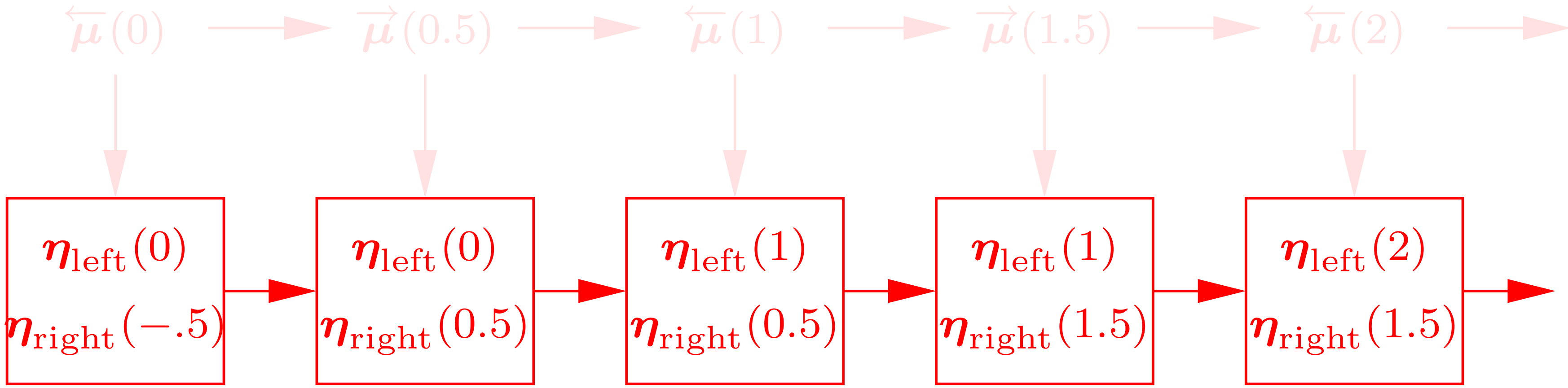
Comment on Macrostates

$$
\begin{tikzcd}
\overleftarrow{\boldsymbol{\mu}}(0) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(0.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(1) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(1.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(2) \arrow[r] \arrow[d] & {} \\
\boxed{\begin{matrix} \boldsymbol{\eta}_{\text{left}}(0) \\ \boldsymbol{\eta}_{\text{right}}(-.5) \end{matrix}} \arrow[r] & \boxed{\begin{matrix} \boldsymbol{\eta}_{\text{left}}(0) \\ \boldsymbol{\eta}_{\text{right}}(0.5) \end{matrix}} \arrow[r] & \boxed{\begin{matrix} \boldsymbol{\eta}_{\text{left}}(1) \\ \boldsymbol{\eta}_{\text{right}}(0.5) \end{matrix}} \arrow[r] & \boxed{\begin{matrix} \boldsymbol{\eta}_{\text{left}}(1) \\ \boldsymbol{\eta}_{\text{right}}(1.5) \end{matrix}} \arrow[r] & \boxed{\begin{matrix} \boldsymbol{\eta}_{\text{left}}(2) \\ \boldsymbol{\eta}_{\text{right}}(1.5) \end{matrix}} \arrow[r] & {}
\end{tikzcd}
.
$$

Comment on Macrostates

$$
\begin{tikzcd}
\overleftarrow{\boldsymbol{\mu}}(0) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(0.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(1) \arrow[r] \arrow[d] & \overrightarrow{\boldsymbol{\mu}}(1.5) \arrow[r] \arrow[d] & \overleftarrow{\boldsymbol{\mu}}(2) \arrow[r] \arrow[d] & {} \\
\boxed{\begin{array}{c}\boldsymbol{\eta}_{\text{left}}(0) \\ \boldsymbol{\eta}_{\text{right}}(-.5)\end{array}} \arrow[r] & \boxed{\begin{array}{c}\boldsymbol{\eta}_{\text{left}}(0) \\ \boldsymbol{\eta}_{\text{right}}(0.5)\end{array}} \arrow[r] & \boxed{\begin{array}{c}\boldsymbol{\eta}_{\text{left}}(1) \\ \boldsymbol{\eta}_{\text{right}}(0.5)\end{array}} \arrow[r] & \boxed{\begin{array}{c}\boldsymbol{\eta}_{\text{left}}(1) \\ \boldsymbol{\eta}_{\text{right}}(1.5)\end{array}} \arrow[r] & \boxed{\begin{array}{c}\boldsymbol{\eta}_{\text{left}}(2) \\ \boldsymbol{\eta}_{\text{right}}(1.5)\end{array}} \arrow[r] & {} .
\end{tikzcd}
$$

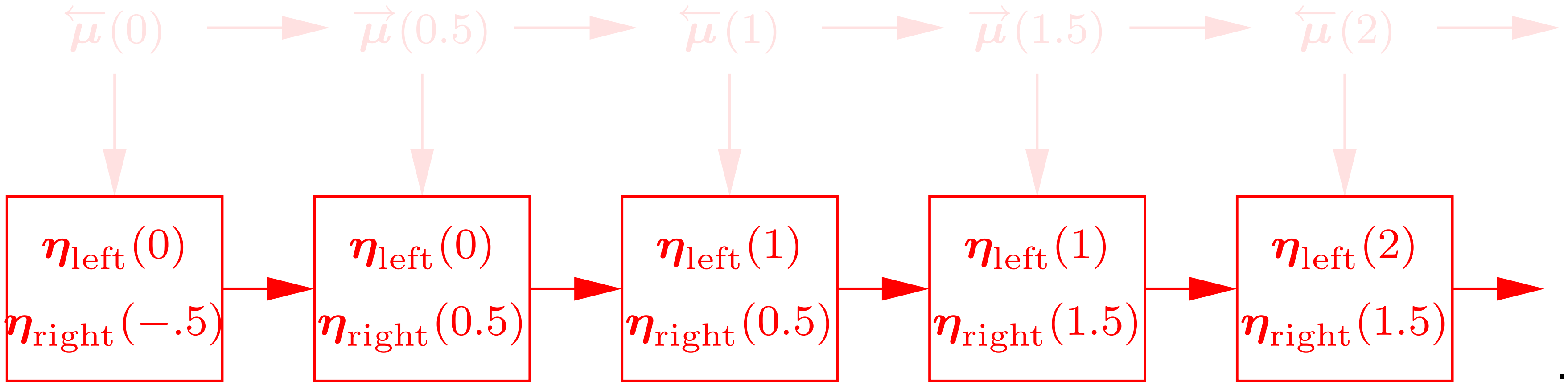
Comment on Macrostates



.

$\Rightarrow$ This can be considered as a "message-free version of the SPA".

Comment on Macrostates



$\Rightarrow$ This can be considered as a "message-free version of the SPA".

Cf. “Message-free version of belief-propagation”
in [Wainwright/Jaakkola/Willsky, 2003].

Bethe Entropy and Weight Spectra

Induced Bethe Entropy

Induced Bethe Entropy

For any $\boldsymbol{\omega} \in \mathcal{P}(\mathbf{H})$, define the induced Bethe entropy to be

$$H_{\text{Bethe}}(\boldsymbol{\omega}) \triangleq H_{\text{Bethe}}(\boldsymbol{\alpha})\Big|_{\boldsymbol{\alpha}=\boldsymbol{\Psi}_{\text{BME}}(\boldsymbol{\omega})} ,$$

where $\boldsymbol{\Psi}_{\text{BME}}(\boldsymbol{\omega})$ is the Bethe max-entropy pseudo-marginal $\boldsymbol{\alpha} \in \mathcal{L}$ among all the pseudo-marginals in $\mathcal{L}$ that correspond to $\boldsymbol{\omega}$.

Evaluating the Bethe Free Entropy Along the Cube Diagonal

Evaluating the Bethe Free Entropy Along the Cube Diagonal

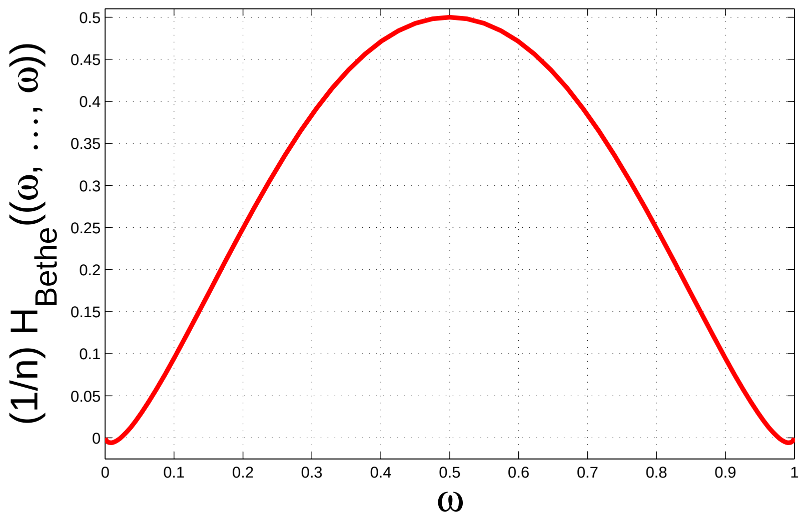
- Take some finite-length (j, k)-regular LDPC code of length n.

Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Take some finite-length (j, k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0, 1]$ we obtain (here for $(j, k) = (3, 6)$):

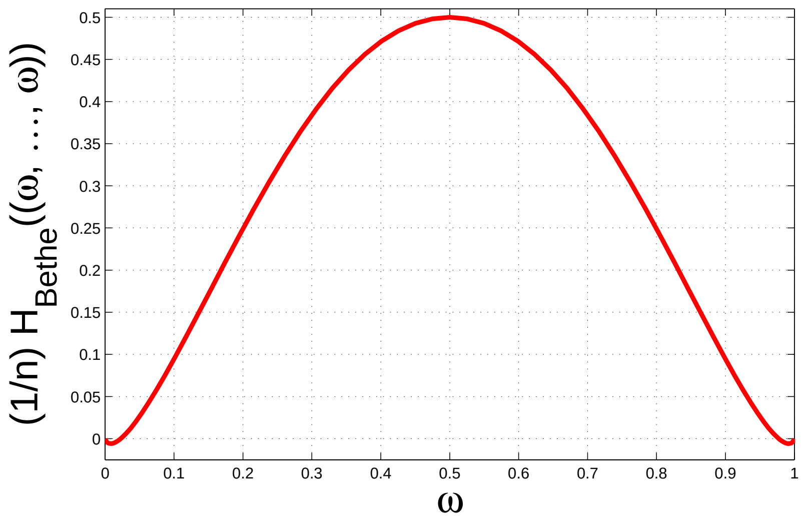
Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Take some finite-length (j, k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0, 1]$ we obtain (here for $(j, k) = (3, 6)$):



Evaluating the Bethe Free Entropy Along the Cube Diagonal

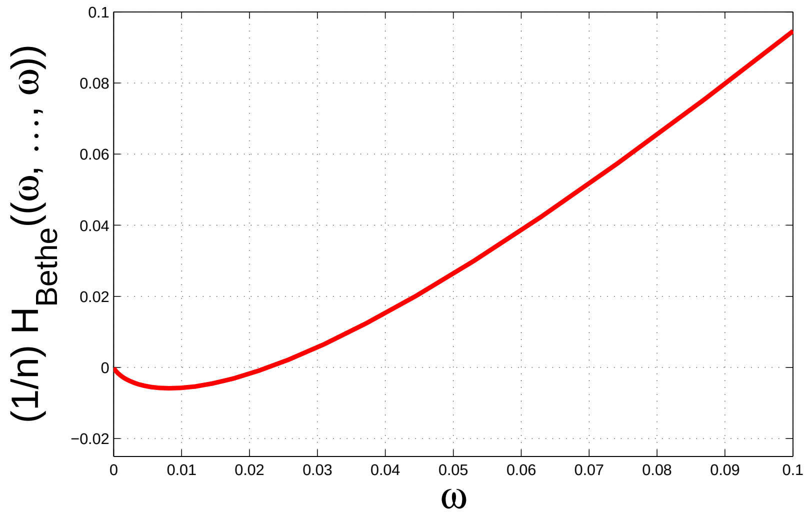
- Take some finite-length (j, k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\text{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0, 1]$ we obtain (here for $(j, k) = (3, 6)$):



- This function happens to equal the exponent of the asymptotic average Hamming weight distribution for Gallager's ensemble of (j, k)-regular LDPC codes!

Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Take some finite-length (j, k)-regular LDPC code of length n.
- Evaluating $\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$ for $\omega \in [0, 1]$ we obtain (here for $(j, k) = (3, 6)$):



- This function happens to equal the exponent of the asymptotic average Hamming weight distribution for Gallager's ensemble of (j, k)-regular LDPC codes!

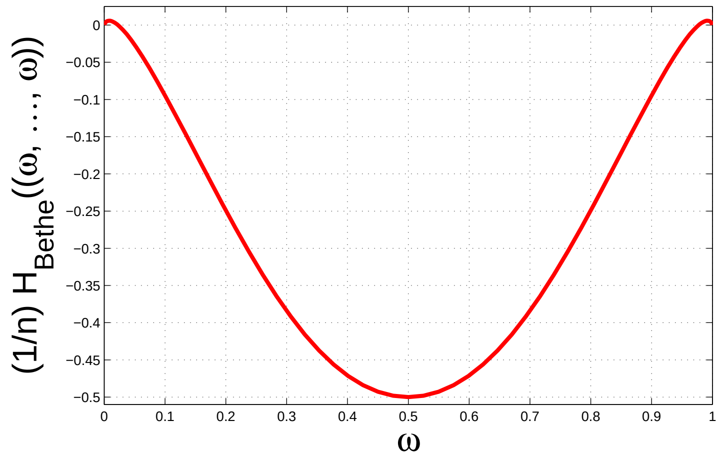
Evaluating the Bethe Free Entropy Along the Cube Diagonal

Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Let's look at $-\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.

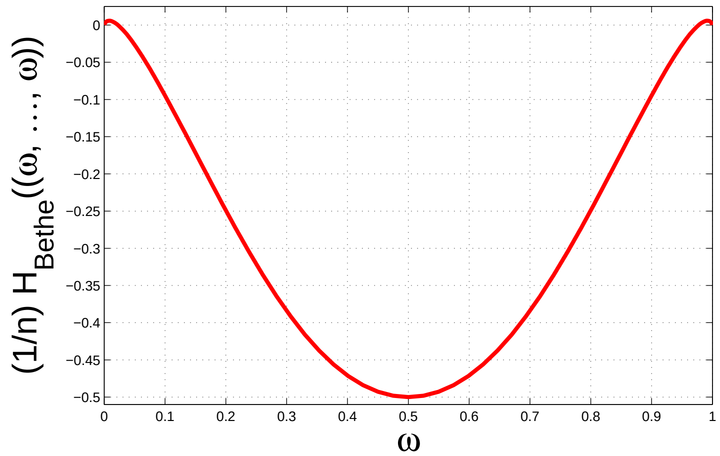
Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Let's look at $-\frac{1}{n}H_{\mathrm{Bethe}}\big((\omega,\ldots,\omega)\big)$.



Evaluating the Bethe Free Entropy Along the Cube Diagonal

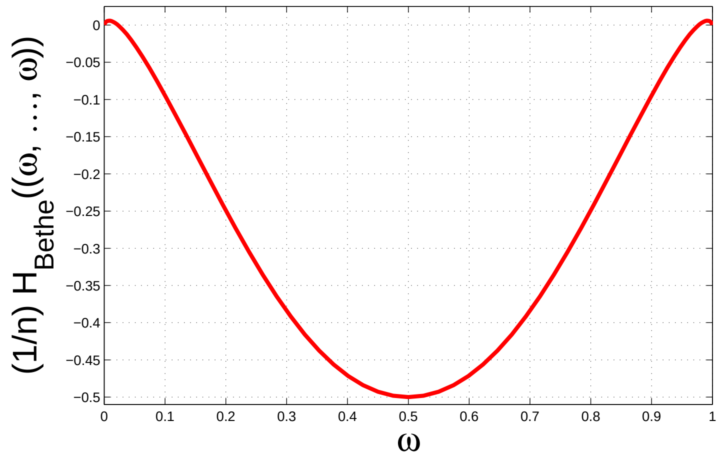
- Let's look at $-\frac{1}{n} H_{\text{Bethe}}\big((\omega, \ldots, \omega)\big)$.



- Remember that $F_{\text{Bethe}}(\boldsymbol{\omega}) = U_{\text{Bethe}}(\boldsymbol{\omega}) - H_{\text{Bethe}}(\boldsymbol{\omega})$.

Evaluating the Bethe Free Entropy Along the Cube Diagonal

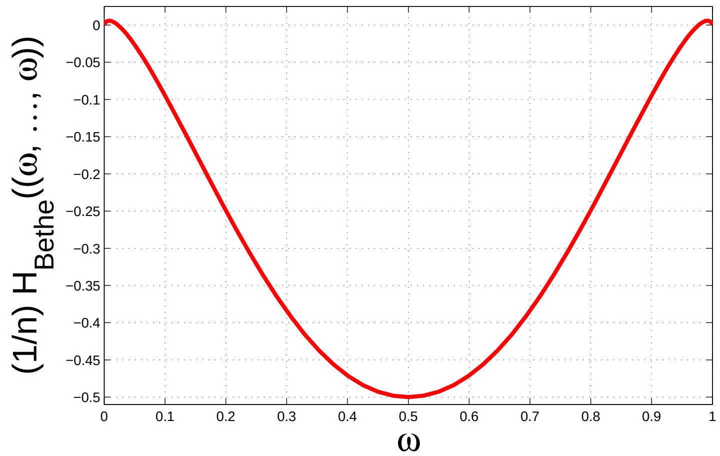
- Let's look at $-\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.



- Remember that $F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega})$.
- Remember that $U_{\mathrm{Bethe}}(\boldsymbol{\omega})$ is linear in $\boldsymbol{\omega}$.

Evaluating the Bethe Free Entropy Along the Cube Diagonal

- Let's look at $-\frac{1}{n} H_{\mathrm{Bethe}}\big((\omega, \ldots, \omega)\big)$.



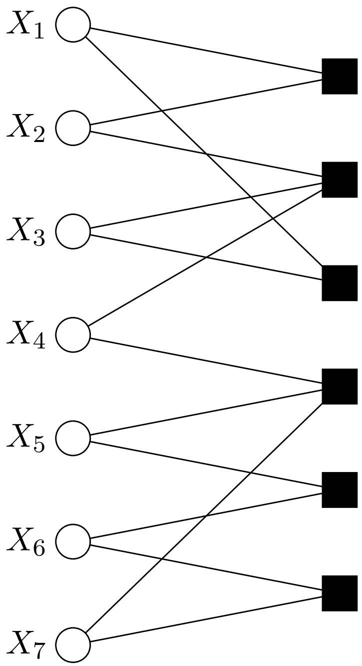
- Remember that $F_{\mathrm{Bethe}}(\boldsymbol{\omega}) = U_{\mathrm{Bethe}}(\boldsymbol{\omega}) - H_{\mathrm{Bethe}}(\boldsymbol{\omega})$.
- Remember that $U_{\mathrm{Bethe}}(\boldsymbol{\omega})$ is linear in $\boldsymbol{\omega}$.
- Therefore, we see that for a finite-length code from an ensemble with asymptotically linearly growing minimum Hamming distance, $F_{\mathrm{Bethe}}(\boldsymbol{\omega})$ is not a convex function of $\boldsymbol{\omega}$.

Bethe Entropy and the Edge Zeta Function

Tanner/Factor Graph of a Cycle Code

Cycle codes are codes which have a Tanner/factor graph where all bit nodes have degree two. (Equivalently, the parity-check matrix has two ones per column.)

Example:

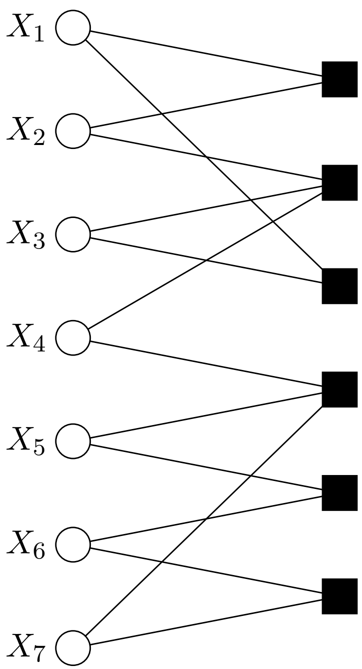


Tanner/factor graph

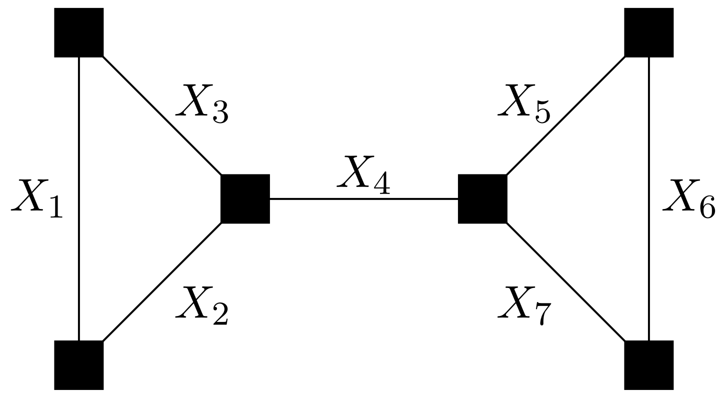
Tanner/Factor Graph of a Cycle Code

Cycle codes are codes which have a Tanner/factor graph where all bit nodes have degree two. (Equivalently, the parity-check matrix has two ones per column.)

Example:



Tanner/factor graph

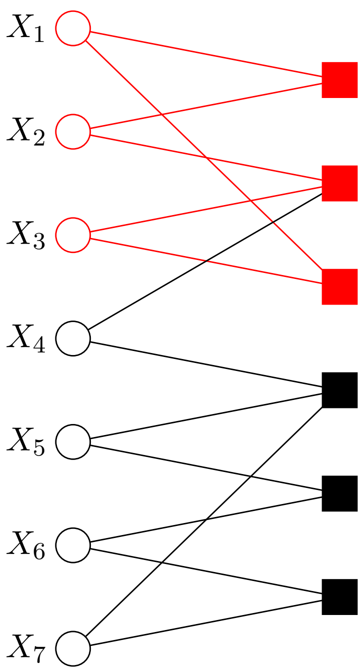


Corresponding normal factor graph

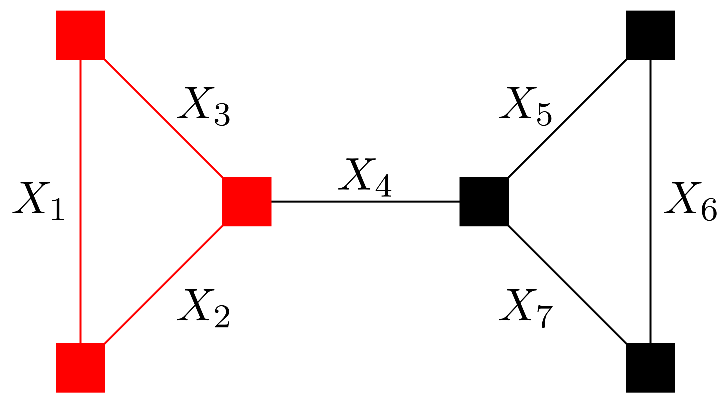
Tanner/Factor Graph of a Cycle Code

Cycle codes are called cycle codes because codewords correspond to simple cycles (or to the symmetric difference set of simple cycles) in the Tanner/factor graph.

Example:



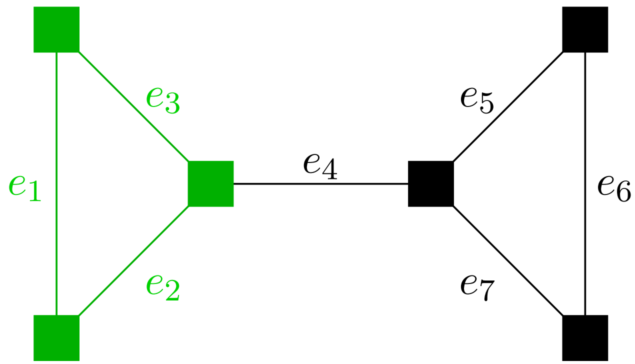
Tanner/factor graph



Corresponding normal factor graph

The Edge Zeta Function of a Graph

Definition (Hashimoto, see also Stark/Terras):



Here: $\Gamma = (e_1, e_2, e_3)$

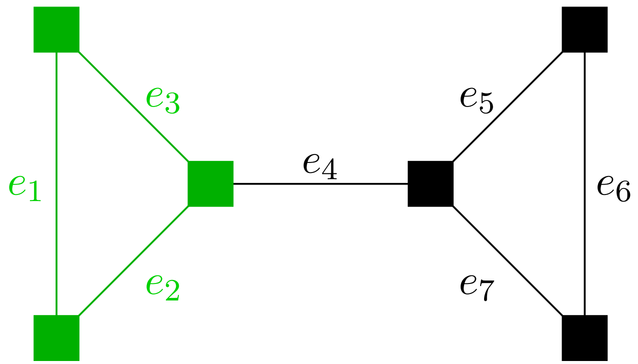
Let Γ be a path in a graph X with edge-set E; write

$$\Gamma = (e_{i_1}, \ldots, e_{i_k})$$

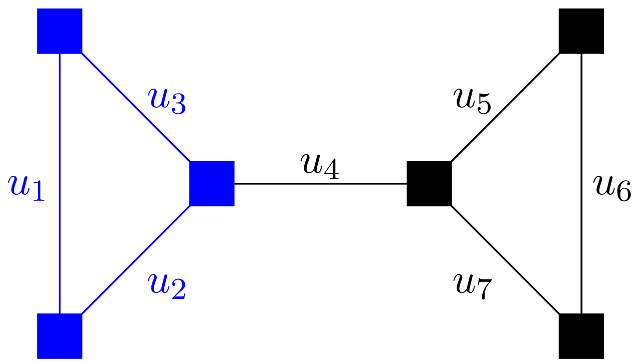
to indicate that Γ begins with the edge e_{i_1} and ends with the edge e_{i_k}.

The Edge Zeta Function of a Graph

Definition (Hashimoto, see also Stark/Terras):



Here: $\Gamma = (e_1, e_2, e_3)$



Here: $g(\Gamma) = u_1 u_2 u_3$

Let Γ be a path in a graph X with edge-set E; write

$$\Gamma = (e_{i_1}, \ldots, e_{i_k})$$

to indicate that Γ begins with the edge e_{i_1} and ends with the edge e_{i_k}.

The monomial of Γ is given by

$$g(\Gamma) \triangleq u_{i_1} \cdots u_{i_k},$$

where the u_i's are indeterminates.

The Edge Zeta Function of a Graph

Definition (Hashimoto, see also Stark/Terras):
The edge zeta function of X is defined to be the power series

$$\zeta_X(u_1, \ldots, u_n) \in \mathbb{Z}[[u_1, \ldots, u_n]]$$

given by

$$\zeta_X(u_1, \ldots, u_n) = \prod_{[\Gamma] \in A(X)} \frac{1}{1 - g(\Gamma)},$$

where $A(X)$ is the collection of equivalence classes of backtrackless, tailless, primitive cycles in X.
Note: unless X contains only one cycle, the set $A(X)$ will be countably infinite.

The Edge Zeta Function of a Graph

Theorem (Bass):

- The edge zeta function $\zeta_X(u_1, \ldots, u_n)$ is a rational function.
- More precisely, for any directed graph $\vec{X}$ of X, we have

$$\zeta_X(u_1, \ldots, u_n) = \frac{1}{\det\left(\mathbf{I} - \mathbf{U}\mathbf{M}(\vec{X})\right)} = \frac{1}{\det\left(\mathbf{I} - \mathbf{M}(\vec{X})\mathbf{U}\right)}$$

where

- $\mathbf{I}$ is the identity matrix of size $2n$,
- $\mathbf{U} = \mathrm{diag}(u_1, \ldots, u_n, u_1, \ldots, u_n)$ is a diagonal matrix of indeterminants.
- $\mathbf{M}(\vec{X})$ is a $2n \times 2n$ matrix derived from some directed graph version $\vec{X}$ of X.

Relationship Pseudo-Codewords and Edge Zeta Function (Part 1: Theorem)

Theorem:

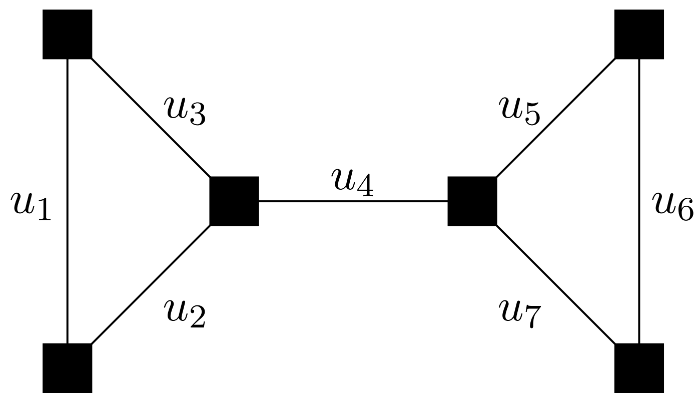
- Let C be a cycle code defined by a parity-check matrix $\mathbf{H}$ having normal graph $N \triangleq N(\mathbf{H})$.
- Let $n = n(N)$ be the number of edges of N.
- Let $\zeta_N(u_1, \ldots, u_n)$ be the edge zeta function of N.
- Then

  the monomial $u_1^{p_1} \ldots u_n^{p_n}$ has a nonzero coefficient
  in the Taylor series expansion of ζ_N

  if and only if

  the corresponding exponent vector $(p_1, \ldots, p_n)$
  is an unscaled pseudo-codeword for C.

Relationship Pseudo-Codewords and Edge Zeta Function (Part 2: Example)

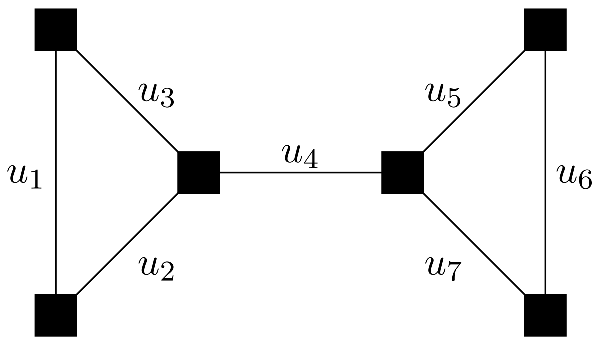


This normal graph N has the following inverse edge zeta function:

$$\zeta_N(u_1, \ldots, u_7) = \frac{1}{\det(\mathbf{I}_{14} - \mathbf{UM})}$$

$$= \frac{1}{\begin{array}{l} 1 - 2u_1u_2u_3 + u_1^2u_2^2u_3^2 - 2u_5u_6u_7 + 4u_1u_2u_3u_5u_6u_7 - 2u_1^2u_2^2u_3^2u_5u_6u_7 \\ -4u_1u_2u_3u_4^2u_5u_6u_7 + 4u_1^2u_2^2u_3^2u_4^2u_5u_6u_7 + u_5^2u_6^2u_7^2 - 2u_1u_2u_3u_5^2u_6^2u_7^2 \\ +u_1^2u_2^2u_3^2u_5^2u_6^2u_7^2 + 4u_1u_2u_3u_4^2u_5^2u_6^2u_7^2 - 4u_1^2u_2^2u_3^2u_4^2u_5^2u_6^2u_7^2 \end{array}}$$

Relationship Pseudo-Codewords and Edge Zeta Function (Part 3: Example)



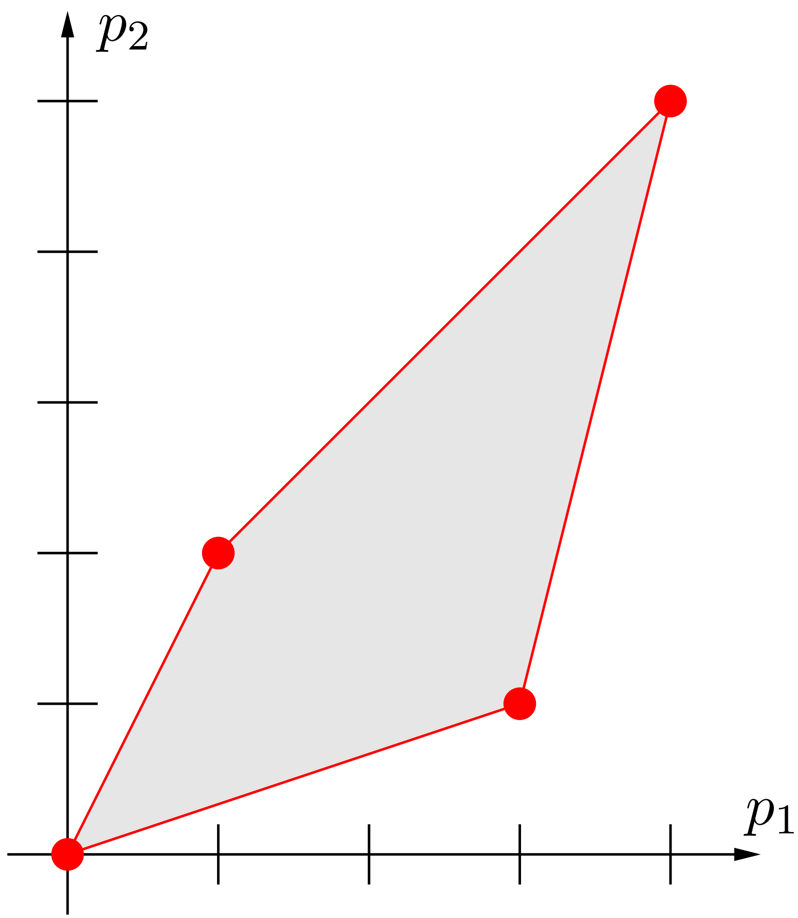
The Taylor series exansion is

$$\zeta_N(u_1, \ldots, u_7)$$
$$= 1 + 2u_1u_2u_3 + 3u_1^2u_2^2u_3^2 + 2u_5u_6u_7$$
$$+ 4u_1u_2u_3u_5u_6u_7 + 6u_1^2u_2^2u_3^2u_5u_6u_7$$
$$+ 4u_1u_2u_3u_4^2u_5u_6u_7 + 12u_1^2u_2^2u_3^2u_4^2u_5u_6u_7$$
$$+ \cdots$$

We get the following exponent vectors:

$(0,0,0,0,0,0,0)$	codeword
$(1,1,1,0,0,0,0)$	codeword
$(2,2,2,0,0,0,0)$	pseudo-codeword (in $\mathbb{Z}$-span)
$(0,0,0,0,1,1,1)$	codeword
$(1,1,1,0,1,1,1)$	codeword
$(2,2,2,0,1,1,1)$	pseudo-codeword (in $\mathbb{Z}$-span)
$(1,1,1,2,1,1,1)$	pseudo-codeword (not in $\mathbb{Z}$-span)
$(2,2,2,2,1,1,1)$	pseudo-codeword (in $\mathbb{Z}$-span)

The Newton Polytope of a Polynomial



Here: $P(u_1, u_2)$

$= u_1^0 u_2^0 + 3u_1^1 u_2^2 + 4u_1^3 u_2^1 - 2u_1^4 u_2^5$

Definition:

The Newton polytope of a polynomial $P(u_1, \ldots, u_n)$ in n indeterminates is the convex hull of the points in n-dimensional space given by the exponent vectors of the nonzero monomials appearing in $P(u_1, \ldots, u_n)$.

Similarly, we can associate a polyhedron to a power series.

Characterizing the Fundamental Cone Through the Zeta Function

Collecting the results from the previous slides we get:

Characterizing the Fundamental Cone Through the Zeta Function

Collecting the results from the previous slides we get:

Proposition: Let $\mathcal{C}$ be some cycle code with parity-check matrix $\mathbf{H}$ and normal factor graph $N(\mathbf{H})$.

The Newton polyhedron of the edge zeta function of $N(\mathbf{H})$

equals

the conic hull of the fundamental polytope $\mathcal{P}(\mathbf{H})$

(aka the fundamental cone $\mathcal{K}(\mathbf{H})$).

Characterizing the Fundamental Cone Through the Zeta Function

Collecting the results from the previous slides we get:

Proposition: Let $\mathcal{C}$ be some cycle code with parity-check matrix $\mathbf{H}$ and normal factor graph $N(\mathbf{H})$.

The Newton polyhedron of the edge zeta function of $N(\mathbf{H})$
equals
the conic hull of the fundamental polytope $\mathcal{P}(\mathbf{H})$
(aka the fundamental cone $\mathcal{K}(\mathbf{H})$).

However, what is the meaning of the coefficients of the monomials in the Taylor series expansion of the edge zeta function?

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Theorem: Let $\boldsymbol{\omega}$ be a pseudo-codeword with rational components. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{\mathrm{Bethe}}(t\boldsymbol{\omega})\bigg|_{t \downarrow 0} \quad = \quad G_{\mathrm{coeff}}\big(\zeta_{\mathsf{N}(\mathbf{H})}, \boldsymbol{\omega}\big).$$

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Theorem: Let $\boldsymbol{\omega}$ be a pseudo-codeword with rational components. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{\mathrm{Bethe}}(t\boldsymbol{\omega})\bigg|_{t\downarrow 0} = G_{\mathrm{coeff}}\big(\zeta_{\mathsf{N}(\mathbf{H})}, \boldsymbol{\omega}\big).$$

- The first term is the directional derivative of the induced Bethe entropy at the origin in the direction of $\boldsymbol{\omega}$.

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Theorem: Let $\boldsymbol{\omega}$ be a pseudo-codeword with rational components. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{\mathrm{Bethe}}(t\boldsymbol{\omega})\bigg|_{t\downarrow 0} \quad = \quad G_{\mathrm{coeff}}\big(\zeta_{\mathsf{N}(\mathbf{H})}, \boldsymbol{\omega}\big).$$

- The first term is the directional derivative of the induced Bethe entropy at the origin in the direction of $\boldsymbol{\omega}$.
- The second term is the growth rate of the coefficients of the monomials that appear in the Taylor series expansion of the edge zeta function $\zeta_{\mathsf{N}(\mathbf{H})}$ and whose exponent vector equals a positive multiple of $\boldsymbol{\omega}$.

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

In the above theorem we used the following definition:

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

In the above theorem we used the following definition:

For every $\boldsymbol{\omega} \in \mathbb{R}_{+}^{|\mathcal{E}|}$ we define

$$G_{\mathrm{coeff}}\big(\zeta_{\mathsf{G}}, \boldsymbol{\omega}\big) \triangleq \limsup_{s \to \infty} \frac{1}{s} \log \operatorname{coeff}\big(\zeta_{\mathsf{G}}(\mathbf{u}), \mathbf{u}^{s \cdot \boldsymbol{\omega}}\big)$$

to be the asymptotic growth rate of the coefficients of the monomials that appear in the Taylor series expansion of the edge zeta function ζ_{G} of the graph G and whose exponent vector equals a positive multiple of $\boldsymbol{\omega}$.

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

In the above theorem we used the following definition:

For every $\boldsymbol{\omega} \in \mathbb{R}_{+}^{|\mathcal{E}|}$ we define

$$G_{\text{coeff}}\big(\zeta_{\mathsf{G}}, \boldsymbol{\omega}\big) \triangleq \limsup_{s \to \infty} \frac{1}{s} \log \operatorname{coeff}\big(\zeta_{\mathsf{G}}(\mathbf{u}), \mathbf{u}^{s \cdot \boldsymbol{\omega}}\big)$$

to be the asymptotic growth rate of the coefficients of the monomials that appear in the Taylor series expansion of the edge zeta function ζ_{G} of the graph G and whose exponent vector equals a positive multiple of $\boldsymbol{\omega}$.

For example, if $\boldsymbol{\omega} = \frac{1}{2} \cdot (1, 1, 1, 2, 1, 1, 1)$ then we consider the asymptotic growth rate of the coefficients of the monomials

$u_1^1 u_2^1 u_3^1 u_4^2 u_5^1 u_6^1 u_7^1, \;\; u_1^2 u_2^2 u_3^2 u_4^4 u_5^2 u_6^2 u_7^2, \;\; u_1^3 u_2^3 u_3^3 u_4^6 u_5^3 u_6^3 u_7^3, \;\; \ldots$

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Theorem: Let $\boldsymbol{\omega}$ be a pseudo-codeword with rational components. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{\mathrm{Bethe}}(t\boldsymbol{\omega})\Big|_{t \downarrow 0} \quad = \quad G_{\mathrm{coeff}}\big(\zeta_{\mathsf{N}(\mathbf{H})}, \boldsymbol{\omega}\big).$$

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Theorem: Let $\boldsymbol{\omega}$ be a pseudo-codeword with rational components. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{\mathrm{Bethe}}(t\boldsymbol{\omega})\bigg|_{t \downarrow 0} \quad = \quad \|\boldsymbol{\omega}\|_1 \cdot H_{\mathrm{MP}}(\boldsymbol{\omega}) \quad = \quad G_{\mathrm{coeff}}\big(\zeta_{\mathsf{N}(\mathbf{H})}, \boldsymbol{\omega}\big).$$

Connecting the Bethe Entropy and the Edge Zeta Function of a Cycle Code

Theorem: Let ω be a pseudo-codeword with rational components. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} H_{\mathrm{Bethe}}(t\boldsymbol{\omega})\bigg|_{t\downarrow 0} \quad = \quad \|\boldsymbol{\omega}\|_1 \cdot H_{\mathrm{MP}}(\boldsymbol{\omega}) \quad = \quad G_{\mathrm{coeff}}\big(\zeta_{\mathsf{N}(\mathbf{H})}, \boldsymbol{\omega}\big).$$

- The first term is the directional derivative of the induced Bethe entropy at the origin in the direction of ω.
- The second term is a scaled version of the entropy rate of some time-invariant Markov process that is associated with ω.
- The third term is the growth rate of the coefficients of the monomials that appear in the Taylor series expansion of the edge zeta function and whose exponent vector equals a positive multiple of ω.

Another Result about the Bethe Entropy around the Origin

Theorem (second-order derivative result of the Bethe entropy):
The larger the eigenvalue gap between the largest and second-largest eigenvalue of the adjacency matrix of the normal factor graph, the larger the curvature of the Bethe entropy around the origin.

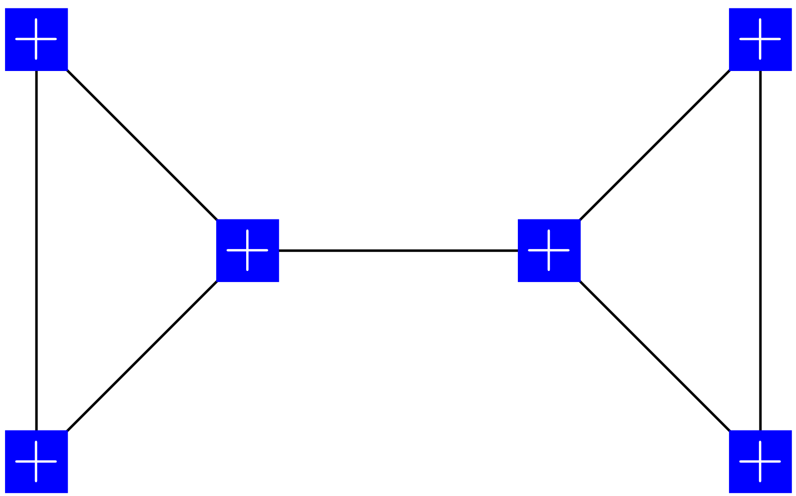
Another Result about the Bethe Entropy around the Origin

Theorem (second-order derivative result of the Bethe entropy):
The larger the eigenvalue gap between the largest and second-largest eigenvalue of the adjacency matrix of the normal factor graph, the larger the curvature of the Bethe entropy around the origin.

$\Rightarrow$ Use so-called Ramanujan graphs to obtain graphical models whose Bethe entropy has maximal curvature around the origin.

Connection to Results by Watanabe and Fukumizu

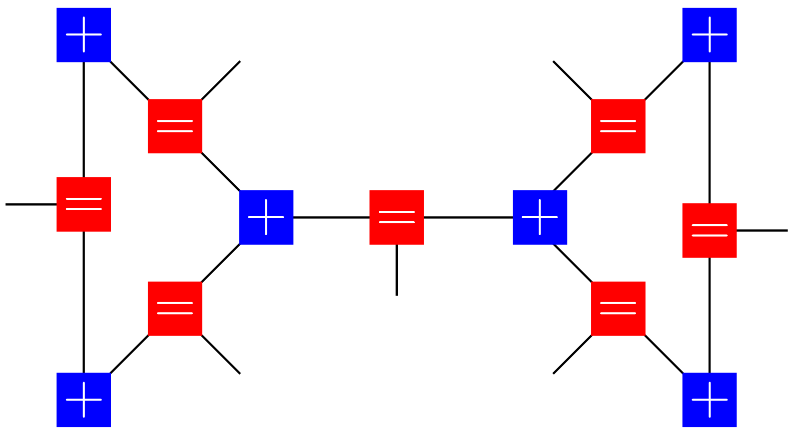
It turns out that the our setup and the setup by [Watanabe/Fukumizu, this workshop / NIPS 2009] are dual to each other in the sense that the sets of valid configurations are given by dual Forney-style factor graphs.



our setup

Connection to Results by Watanabe and Fukumizu

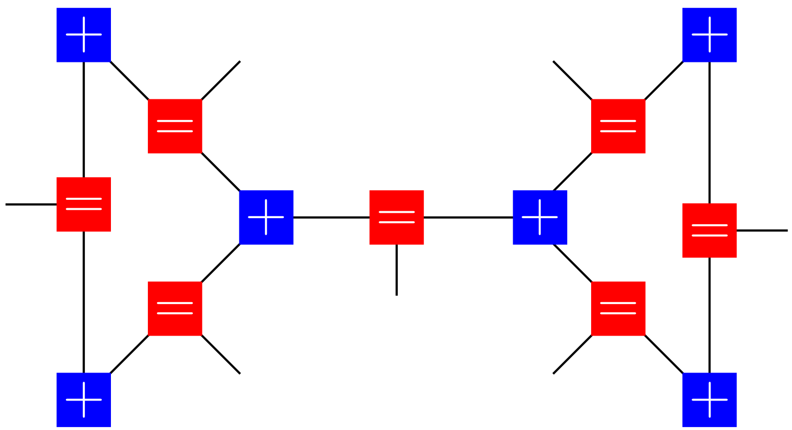
It turns out that the our setup and the setup by [Watanabe/Fukumizu, this workshop / NIPS 2009] are dual to each other in the sense that the sets of valid configurations are given by dual Forney-style factor graphs.



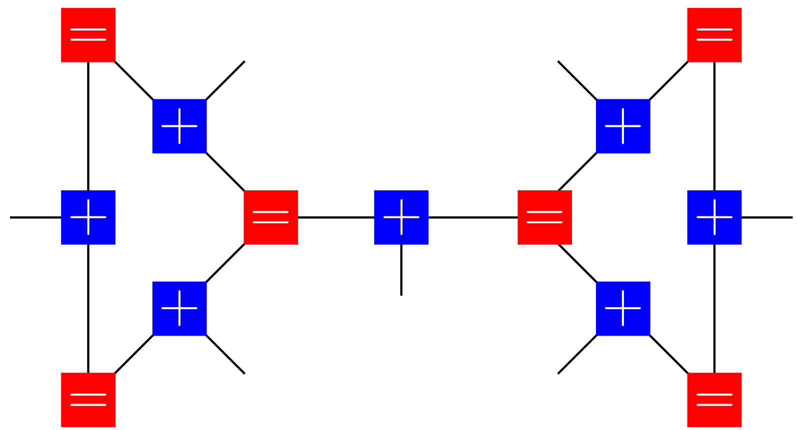
our setup

Connection to Results by Watanabe and Fukumizu

It turns out that the our setup and the setup by [Watanabe/Fukumizu, this workshop / NIPS 2009] are dual to each other in the sense that the sets of valid configurations are given by dual Forney-style factor graphs.



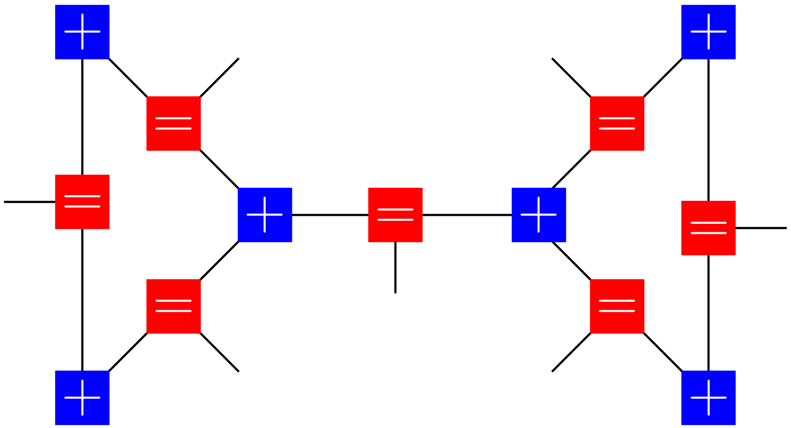
our setup



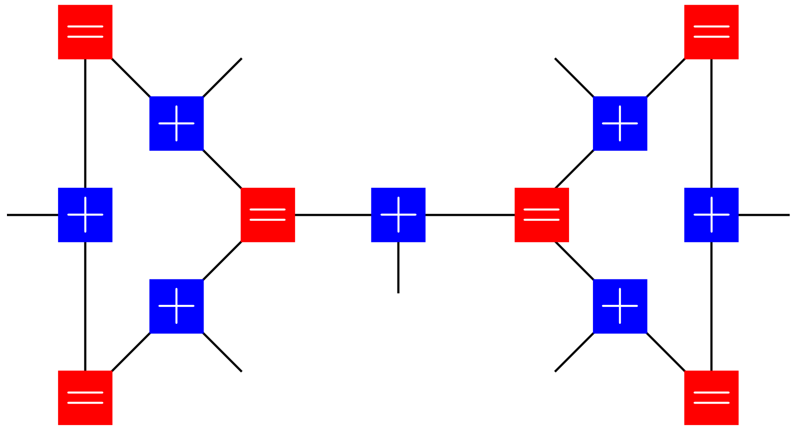
Watanabe/Fukumizu setup

Connection to Results by Watanabe and Fukumizu

It turns out that the our setup and the setup by [Watanabe/Fukumizu, this workshop / NIPS 2009] are dual to each other in the sense that the sets of valid configurations are given by dual Forney-style factor graphs.



our setup



Watanabe/Fukumizu setup

However, the type of obtained results are quite different.

Conclusions / Outlook

Conclusions / Outlook

- We have given an interpretation of the Bethe entropy and a re-interpretation of the theorem by Yedidia/Freeman/Weiss.

Conclusions / Outlook

- We have given an interpretation of the Bethe entropy and a re-interpretation of the theorem by Yedidia/Freeman/Weiss.

  $\Rightarrow$ A similar (re-)interpretation can also be given for the Kikuchi entropy and generalized belief propagation.

Conclusions / Outlook

- We have given an interpretation of the Bethe entropy and a re-interpretation of the theorem by Yedidia/Freeman/Weiss.

  $\Rightarrow$ A similar (re-)interpretation can also be given for the Kikuchi entropy and generalized belief propagation.

  $\Rightarrow$ Also new insights also for survey propagation?

Conclusions / Outlook

- We have given an interpretation of the Bethe entropy and a re-interpretation of the theorem by Yedidia/Freeman/Weiss.

  $\Rightarrow$ A similar (re-)interpretation can also be given for the Kikuchi entropy and generalized belief propagation.

  $\Rightarrow$ Also new insights also for survey propagation?

- We have discussed why from a coding-theoretic point of view it can be desirable to have a convex-concave Bethe variation free energy.

Conclusions / Outlook

- We have given an interpretation of the Bethe entropy and a re-interpretation of the theorem by Yedidia/Freeman/Weiss.
  - $\Rightarrow$ A similar (re-)interpretation can also be given for the Kikuchi entropy and generalized belief propagation.
  - $\Rightarrow$ Also new insights also for survey propagation?
- We have discussed why from a coding-theoretic point of view it can be desirable to have a convex-concave Bethe variation free energy.
- We have discussed a connection between Bethe entropy and the edge zeta function of cycle codes.

Conclusions / Outlook

- We have given an interpretation of the Bethe entropy and a re-interpretation of the theorem by Yedidia/Freeman/Weiss.

  $\Rightarrow$ A similar (re-)interpretation can also be given for the Kikuchi entropy and generalized belief propagation.

  $\Rightarrow$ Also new insights also for survey propagation?

- We have discussed why from a coding-theoretic point of view it can be desirable to have a convex-concave Bethe variation free energy.

- We have discussed a connection between Bethe entropy and the edge zeta function of cycle codes.

  $\Rightarrow$ See also the talk on Friday morning by Watanabe.


Thank you!
www.pseudocodewords.info

LABS hp